



Characteristic function-based hypothesis tests under weak dependence

Anne Leucht

Friedrich-Schiller-Universität Jena, Institut für Stochastik, Ernst-Abbe-Platz 2, D-07743 Jena, Germany

ARTICLE INFO

Article history:

Received 17 May 2011

Available online 18 February 2012

AMS 2010 subject classifications:

62G10

62E20

62F40

Keywords:

Bootstrap

Empirical characteristic function

Goodness-of-fit

Symmetry test

Time series

V-statistics

ABSTRACT

In this article we propose two consistent hypothesis tests of L_2 -type for weakly dependent observations based on the empirical characteristic function. We consider a symmetry test and a goodness-of-fit test for the marginal distribution of a time series. The asymptotic behaviour under the null as well as fixed and certain local alternatives is investigated. Since the limit distributions of the test statistics depend on unknown parameters in a complicated way, we suggest to apply certain parametric bootstrap methods in order to determine critical values of the tests.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction and motivation

Time series models are well-established tools to describe various real-life phenomena, e.g. in finance and biometrics. In the present paper we provide consistent testing procedures that allow for checking adequacy of a certain model with respect to symmetry and parametric structure of the marginal distribution of a time series.

Let X_1, \dots, X_n be \mathbb{R}^d -valued observations of a stationary weakly dependent process with marginal distribution P_X . We consider problems of the following structure:

$$\mathcal{H}_0 : g_X(\cdot) = f_X(\cdot, \theta_0) \quad \text{for some } \theta_0 \in \Theta \subseteq \mathbb{R}^p \text{ against}$$

$$\mathcal{H}_1 : g_X(\cdot) \neq f_X(\cdot, \theta) \quad \forall \theta \in \Theta.$$

Besides supremum-type tests, L_2 -tests are the most convenient ones in mathematical statistics. We apply a test statistic of L_2 -type, $\hat{T}_n = n \int_{\mathbb{R}^d} |\psi(c_n)(t) - f_X(t, \hat{\theta}_n)|^2 w(t) dt$. Here, $\hat{\theta}_n$ is a consistent estimator of θ_0 and c_n denotes the empirical characteristic function, i.e. $c_n(t) = n^{-1} \sum_{k=1}^n e^{it'X_k}$. The function ψ is chosen such that $\psi(c_n)$ forms an appropriate nonparametric estimate of $g_X = \psi(c_X)$, where c_X is the characteristic function of X_1 . We specify the functions ψ and f_X for the concrete test problems in Sections 2 and 3. There are several approaches in order to derive the asymptotic null distribution of \hat{T}_n . Under certain regularity conditions the latter statistic can be approximated by an L_2 -statistic $T_n = \int_{\mathbb{R}^d} |n^{-1/2} \sum_{k=1}^n h(X_k, t, \theta_0)|^2 w(t) dt$ with a certain continuous function h and with fixed parameter θ_0 . The limit of latter quantity can be deduced invoking empirical process theory. Alternatively, $n^{-1/2} \sum_{k=1}^n h(X_k, t, \theta_0)$, $t \in \mathbb{R}^d$, can be viewed as an element of the Hilbert space $L_2(\mathbb{R}^d, \mathcal{B}^d, w(t) dt)$. Therefore, the asymptotics of T_n can also be obtained invoking a

E-mail address: anne.leucht@uni-jena.de.

central limit theorem (CLT) for random variables with values in a Hilbert space and the continuous mapping theorem. In contrast, the approximating statistic can also be understood as a degenerate V -statistic. In the present paper, we employ a recent result on degenerate V -statistics by Leucht [25] that allows for the derivation of the limit distributions under easily verifiable assumptions. In order to study the behaviour of our test statistics under local alternatives, this approach is carried over to a special kind of triangular schemes of random variables; cf. Section 4.2.

In Section 2, we extend the symmetry test which was initially proposed by Feuerverger and Mureika [17] for independent and identically distributed (i.i.d.) data and a known centre of symmetry to the case of weakly dependent observations with unknown location parameter. Answering the question whether a distribution is symmetric or not is interesting for several reasons. First, there is a pure statistical interest since presence or absence of symmetry is important for deciding what parameter to estimate. If the underlying distribution is symmetric, the point of symmetry is the only reasonable location parameter, whereas in the non-symmetric case there is no longer only one measure of location; cf. [1]. Moreover, robust estimators of location, e.g. trimmed means, and robust tests for location parameters assume the underlying observations to arise from a symmetric distribution; see for instance [34]. Consequently it is important to check this assumption before applying those methods. Rejecting the hypothesis of symmetry also has substantial impact on model selection. In this case, fitting data to nonlinear AR(p) processes with skew symmetric regression functions and symmetric innovations is inappropriate; cf. [30]. Finally, symmetry plays a central role in analysing and modelling business circles since time reversibility of a process $(X_t)_{t \in \mathbb{Z}}$, i.e. $P_{X_{t_1}, \dots, X_{t_k}} = P_{X_{t_k}, \dots, X_{t_1}}$, $t_1, \dots, t_k \in \mathbb{Z}$, $k \in \mathbb{N}$, implies symmetry of the distributions of the differences $Y_{t,k} = X_t - X_{t-k}$, $k \in \mathbb{N}$, about the origin. For a detailed discussion of this topic see [31,10].

A great variety of symmetry tests for i.i.d. random variables are available in the literature. Detailed lists of references are given by Lee [24] as well as by Henze et al. [20]. Also in the context of time series numerous tests of symmetry have been employed. There are several moment-based tests, e.g. by Ramsey and Rothman [31] or Bai and Ng [3]. However, both approaches are inconsistent against alternatives whose third moments vanish but still are not symmetric. Lee [24] built an M test for the composite hypothesis of symmetry around an unknown location parameter and derived asymptotic normality under the null and certain local alternatives. However, a consistency result is not stated. Fan and Ullah [16] considered a consistent L_2 -test for the simple hypothesis based on kernel density estimates with vanishing bandwidth under the assumption that the sample arises from a certain absolute regular process. The test for the simple hypothesis by Chen et al. [10] employs the fact that a distribution is symmetric if and only if the imaginary part of its characteristic function $\Im c(t) = \mathbb{E} \sin(t'X)$ vanishes. No moment restrictions on the marginal distributions of the underlying data are required. They investigated the asymptotics under the null and certain local alternatives but they did not derive consistency of their test. We also establish a characteristic function-based test for symmetry that is asymptotically unbiased against Pitman local alternatives. In contrast to [10], we consider the composite hypothesis of symmetry around an unknown centre and obtain consistency of our approach.

In Section 3, a goodness-of-fit test for the marginal distribution of a time series is constructed. So far, the normality assumption is still dominating the literature in many fields, however, not for reasons of empirical evidence but for theoretical simplicity. Within the last decades the literature on non-Gaussian time series has addressed the topic of finding distributions of the innovations corresponding to specified marginals of a certain model. Surveys of those results are given by Jose et al. [23] as well as by Block et al. [7].

While there is a great variety of tests concerning the problem whether the distribution of a sample belongs to a parametric class of distributions if the underlying observations are i.i.d., the number of consistent tests is limited in the dependent case. Recently, Ignaccolo [21] and Munk et al. [26] developed results for α -mixing variables generalizing Neyman's smooth test for a simple null hypothesis. Tests for the composite hypothesis based on the L_2 -difference between a smoothed version of the parametric density estimate and a nonparametric estimator were considered by Fan and Ullah [16] as well as by Neumann and Paparoditis [28]. The bandwidths involved in the estimators were supposed to be asymptotically vanishing. A drawback resulting from the latter assumption is the loss of power against Pitman local alternatives compared to approaches with fixed bandwidths; see [18]. Motivated by the fact that tests based on kernel density estimators with decreasing smoothing parameter are very sensitive to the choice of the bandwidth, Fan [15] established a characteristic-function based test for the composite hypothesis in the i.i.d. case that evaluates the difference between the characteristic function and its fixed-kernel estimate, the empirical characteristic function. To the author's best knowledge this approach has not been considered if the observations are dependent. Here, we extend the L_2 -test of [15] to the time series setting. This type of test is useful in particular when the density has a complicated structure while the characteristic function is of simple form. A typical example of such a distribution is the normal inverse Gaussian distribution that is widely used in mathematical finance, see e.g. [4].

For the above-named tests, we derive their limit distributions under the respective null hypotheses and under certain local alternatives. It turns out that already in the case of i.i.d. observations the asymptotic distributions depend on unknown parameters in a complicated way. The bootstrap offers a convenient way to determine critical values of the tests. However, a naive application of these resampling methods fails. We obtain bootstrap consistency after a modification of the test statistics on the bootstrap side. It is well-known that model-based bootstrap counterparts of the underlying observations can often not be proved to be mixing even though the original process satisfies some mixing condition. However, other concepts of weak dependence can be verified, see e.g. [5] or [14]. Throughout the paper, the underlying process is assumed to meet the following definition, which is due to [12].

Definition 1.1. Let (Ω, \mathcal{A}, P) be a probability space and $(X_k)_{k \in \mathbb{N}}$ be a stationary sequence of integrable \mathbb{R}^d -valued random variables. The process is called τ -dependent if

$$\tau_r = \sup_{l \in \mathbb{N}} \frac{1}{l} \sup_{r+i \leq k_1, \dots, k_l} \{\tau(\sigma(X_1, \dots, X_i), (X_{k_1}, \dots, X_{k_l}))\} \xrightarrow{r \rightarrow \infty} 0,$$

where

$$\tau(\mathcal{M}, X) = \mathbb{E} \left(\sup_{f \in \Lambda_1(\mathbb{R}^p)} \left| \int_{\mathbb{R}^p} f(x) dP_{X|\mathcal{M}}(x) - \int_{\mathbb{R}^p} f(x) dP_X(x) \right| \right).$$

Here, \mathcal{M} is a sub- σ -algebra of \mathcal{A} , $P_{X|\mathcal{M}}$ denotes the conditional distribution of the \mathbb{R}^p -valued random variable X given \mathcal{M} , and $\Lambda_1(\mathbb{R}^p)$ denotes the set of 1-Lipschitz functions from \mathbb{R}^p to \mathbb{R} .

If Ω is rich enough, there exists a random variable \tilde{X} that is independent of $\mathcal{M} = \sigma(X_1, \dots, X_i)$ such that $\tilde{X} \stackrel{d}{=} X$ and

$$\tau(\mathcal{M}, X) = \mathbb{E} \|X - \tilde{X}\|, \quad (1.1)$$

where $\|x\| = \sum_{j=1}^p |x_j|$, $x \in \mathbb{R}^p$. Actually, $\tau(\mathcal{M}, X)$ is the minimal L_1 -distance between X and any $Y \stackrel{d}{=} X$ that is independent of \mathcal{M} . This L_1 -coupling property of the τ -coefficient was verified by Dedecker and Prieur [13] and will be essential for all our proofs below. Dedecker and Prieur [12] discussed relations of the τ -coefficient to ordinary mixing coefficients. Additionally, they provided an extensive list of examples for τ -dependent processes including causal linear and functional autoregressive processes. Note that also ARMA processes with weakly dependent innovations and GARCH processes are τ -dependent; see [33]. We conclude this section mentioning that all proofs and asymptotic results on V -statistics under local alternatives are deferred to a final Section 4.

2. Testing symmetry

2.1. The test statistic and its asymptotics

Suppose that X_1, \dots, X_n are \mathbb{R}^d -valued random variables with distribution P_X and characteristic function c_X . We establish a test for the problem

$$\mathcal{H}_0: P_{X-\mu_0} = P_{\mu_0-X} \quad \text{for some } \mu_0 \in \mathbb{R}^d \quad \text{vs.} \quad \mathcal{H}_1: P_{X-\mu} \neq P_{\mu-X} \quad \forall \mu \in \mathbb{R}^d$$

that can be reformulated in terms of the imaginary part of the corresponding characteristic functions:

$$\mathcal{H}_0: \Im(c_{X-\mu_0}) \equiv 0 \quad \text{for some } \mu_0 \in \mathbb{R}^d \quad \text{vs.} \quad \mathcal{H}_1: \Im(c_{X-\mu}) \neq 0 \quad \forall \mu \in \mathbb{R}^d.$$

Inspired by Feuerverger and Mureika [17], who tested for symmetry around the origin in the case of i.i.d. univariate observations, and Henze et al. [20], who investigated the composite hypothesis for i.i.d. multivariate random variables, we suggest the test statistic

$$\hat{S}_n = n \int_{\mathbb{R}^d} [\Im(e^{-it'\hat{\mu}_n} c_n(t))]^2 w(t) dt = \int_{\mathbb{R}^d} \left[\frac{1}{\sqrt{n}} \sum_{k=1}^n \sin(t'(X_k - \hat{\mu}_n)) \right]^2 w(t) dt.$$

Here, $\hat{\mu}_n$ denotes an estimator of μ_0 . We make the following assumptions:

(S1)(i) $(X_n)_{n \in \mathbb{N}}$ is a sequence of τ -dependent random variables with values in \mathbb{R}^d and with $\sum_{r=1}^{\infty} r(\tau_r)^{\delta^2} < \infty$ for some $\delta \in (0, 1)$.

(ii) $\mathbb{E} \|X_1\|^{1+\varepsilon} < \infty$ and all $\varepsilon \in (0, 1/(3-2\delta))$.

(iii) The sequence of estimators $\hat{\mu}_n$ admits the expansion

$$\hat{\mu}_n - \mu_0 = \frac{1}{n} \sum_{k=1}^n l(X_k, \mu_0) + o_p(n^{-1/2}),$$

where $l: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a Lipschitz continuous function satisfying $\mathbb{E} l(X_1, \mu_0) = 0_d$ and $\mathbb{E} \|l(X_1, \mu_0)\|^{2\nu} < \infty$ for some $\nu > (2-\delta)/(1-\delta)$. (Here, 0_d denotes the d -dimensional null vector.)

(iv) The function w is measurable, positive almost everywhere w.r.t. the Lebesgue measure on \mathbb{R}^d and satisfies $\int_{\mathbb{R}^d} (1 + \|t\|^{2\nu}) w(t) dt < \infty$.

Remark 2.1. (i) The assumption of l being Lipschitz continuous can be weakened at the expense of additional moment constraints. However, if one simply uses the arithmetic mean to estimate μ_0 , (S1)(ii) is satisfied if $\mathbb{E} \|X_1\|^{2\nu} < \infty$.

- (ii) Note that it does not suffice to postulate strict positivity of w on some interval around the centre of symmetry since there are non-symmetric distributions whose characteristic functions are real in a certain neighbourhood of the origin. For an example, see [36, Ex. 21, Appendix A].

Under these weak constraints \widehat{S}_n can be approximated by a degenerate V -statistic (multiplied by n) with fixed parameter μ_0 ,

$$S_n = \frac{1}{n} \sum_{j,k=1}^n \int_{\mathbb{R}^d} [\sin(t'(X_j - \mu_0)) - c_{X-\mu_0}(t)t'l(X_j, \mu_0)][\sin(t'(X_k - \mu_0)) - c_{X-\mu_0}(t)t'l(X_k, \mu_0)]w(t) dt.$$

Lemma 2.1. Suppose that assumption (S1) holds. Then, under \mathcal{H}_0 ,

$$\widehat{S}_n - S_n = o_P(1).$$

Due to the above approximation, the asymptotic distributions of the test statistic and the degenerate V -statistic S_n coincide. We obtain the limit of the latter variable by applying Theorem 2.2 of [25].

Theorem 2.1. Suppose that assumption (S1) holds. Then, under the null hypothesis, there are families of constants $(\gamma_{j;k_1,k_2}^{(c)}(\mu_0))_{c,j,k_1,k_2}$, $(C_{j;k_1,k_2})_{j,k_1,k_2}$ and centred, jointly normal random variables $(Z_{j;k})_{j,k}$ such that

$$\begin{aligned} \widehat{S}_n \xrightarrow{d} Z_1 &:= \lim_{c \rightarrow \infty} \sum_{j=1}^{\infty} \sum_{k_1, k_2 \in \mathbb{Z}^d} \gamma_{j;k_1,k_2}^{(c)}(\mu_0) [Z_{j;k_1} Z_{j;k_2} + C_{j;k_1,k_2}] \\ &+ \mathbb{E} \int_{\mathbb{R}^d} [\sin(t'(X_1 - \mu_0)) - c_{X-\mu_0}(t)t'l(X_1, \mu_0)]^2 w(t) dt. \end{aligned}$$

Here, the r.h.s. converges in the L_2 -sense.

Remark 2.2. (i) The limit distribution of the test statistic depends on the unknown parameter and the underlying dependence structure in a complicated way. Therefore, problems arise as soon as (asymptotic) critical values of the test have to be determined. In the following subsection we propose the application of certain bootstrap methods in order to circumvent these difficulties.

- (ii) The approximating statistic can be rewritten as $S_n = \|n^{-1/2} \sum_{k=1}^n Y_k\|_{L_2(\mathbb{R}^d, w(t)dt)}$ with $Y_k(t) = \sin(t'(X_k - \mu_0)) - c_{X-\mu_0}(t)t'l(X_k, \mu_0)$, $t \in \mathbb{R}^d$. Therefore, its limit distribution could also be derived using a CLT for Hilbert space valued random variables and the continuous mapping theorem. For example, Neuhaus [27] and Dedecker and Merlevède [11] employed this approach to derive the asymptotics of the ordinary Cramér–von Mises statistic in the i.i.d. and weakly dependent case, respectively. The latter method yields a more comprehensive representation of the limit as an infinite weighted sum of i.i.d. $\mathcal{N}(0, 1)$ variables. Still, the weights depend on the unknown parameter and the unknown dependence structure in a complicated way and therefore the analytic calculation of quantiles is impossible. In order to verify the validity of model-based bootstrap for the approximation of critical values of the test, one requires a CLT for triangular schemes of Hilbert space valued random variables. To the author's best knowledge, there is no such result under weak dependence with exception of mixing. However, model-based bootstrap yields samples which can often not be proved to be mixing. For that reason and since an explicit knowledge of the limit is not necessary to apply the bootstrap-based test proposed in the following subsection, we do not pursue the Hilbert space approach to the asymptotics here.

Next, the behaviour of the test statistic under fixed alternatives is considered.

Lemma 2.2. Suppose that (S1)(i) and (iii) hold. Moreover, assume that there exists a constant $\mu_0 \in \mathbb{R}^d$ such that $\widehat{\mu}_n \xrightarrow{P} \mu_0$. Then, under \mathcal{H}_1 ,

$$P(\widehat{S}_n > K) \xrightarrow{n \rightarrow \infty} 1, \quad \forall K < \infty.$$

Finally, we study the asymptotic behaviour of \widehat{S}_n under Pitman local alternatives. More precisely, we consider alternatives

$$\mathcal{H}_{1,n}: \frac{dP_{X_{n,1}}}{dP_{X_1}} = 1 + \frac{g_n}{\sqrt{n}} \quad \text{with } \|g_n - g\|_{\infty} \xrightarrow{n \rightarrow \infty} 0 \quad \text{and} \quad \Im c_{X_{n,1}-\mu}(t) \neq 0, \quad \forall \mu \in \mathbb{R}^d,$$

where g is supposed to be a measurable bounded function. To this end, we assume

- (S2)(i)** The triangular scheme $(X_{n,k})_{k=1}^n$, $n \in \mathbb{N}$, of row-wise stationary and \mathbb{R}^d -valued random variables fulfils $\sum_{r=1}^{\infty} \tau_r(\bar{\tau}_r)^{\delta^2} < \infty$ for some $\delta \in (0, 1)$. Here, the sequence $(\bar{\tau}_r)_{r \in \mathbb{N}}$ is defined as $\bar{\tau}_r := \sup_{n \geq r} \tau_{r,n}$ with

$$\tau_{r,n} := \sup_{l \in \mathbb{N}} \frac{1}{l} \sup_{r+i \leq k_1, \dots, k_l} \{\tau(\sigma(X_{n,1}, \dots, X_{n,i}), (X_{n,k_1}, \dots, X_{n,k_l}))\}.$$

- (ii) There exists a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ satisfying assumption (S1)(i) and such that $(X'_{n,t_1}, X'_{n,t_2})' \xrightarrow{d} (X'_{t_1}, X'_{t_2})'$, $1 \leq t_1, t_2 \leq n$, $n \in \mathbb{N}$. The Radon Nikodym derivatives are given by $dP_{X_{n,1}}/dP_{X_1} = 1 + n^{-1/2}g_n$, where $(g_n)_{n \in \mathbb{N}}$ is a sequence of functions with $\|g_n - g\|_\infty \rightarrow_{n \rightarrow \infty} 0$ for some bounded measurable function g .

In Section 4.2 we establish asymptotic results on V -statistics under local alternatives for dependent random variables. Invoking these findings, we obtain the subsequent assertion.

Proposition 2.1. Suppose that the conditions (S1) and (S2) are satisfied. Additionally, assume that $\hat{\mu}_n(X_{n,1}, \dots, X_{n,n}) - \mu_0 = n^{-1} \sum_{k=1}^n l(X_{n,k}, \mu_0) + o_p(n^{-1/2})$. Then, under $\mathcal{H}_{1,n}$,

$$\begin{aligned} \hat{S}_n \xrightarrow{d} Z_{1,loc} &:= \lim_{c \rightarrow \infty} \sum_{j=1}^{\infty} \sum_{k_1, k_2 \in \mathbb{Z}^d} \gamma_{j; k_1, k_2}^{(c)}(\mu_0) [\{Z_{j; k_1} + D_{j; k_1}\} \{Z_{j; k_2} + D_{j; k_2}\} + C_{j; k_1, k_2}] \\ &\quad + \mathbb{E} \int_{\mathbb{R}^d} [\sin(t'(X_1 - \mu_0)) - c_{X-\mu_0}(t) t' l(X_1, \mu_0)]^2 w(t) dt \end{aligned}$$

with a family of constants $(D_{j;k})_{j,k}$. If additionally $\text{var}(Z_1) > 0$, then

$$\liminf_{n \rightarrow \infty} [P_{\mathcal{H}_{1,n}}(\hat{S}_n > x) - P_{\mathcal{H}_0}(\hat{S}_n > x)] \geq 0, \quad \forall x \in \mathbb{R}.$$

2.2. Bootstrap validity

The foregoing results suggest to reject the hypothesis of symmetry at asymptotic significance level α if $\hat{S}_n > t_\alpha$. Here, t_α denotes the $(1 - \alpha)$ -quantile of the distribution of Z_1 . The latter random variable depends on the unknown parameter and the dependence structure of the underlying process $(X_t)_t$ in a complicated way and thus (asymptotic) critical values can hardly be derived analytically or be tabulated. Therefore, we consider a parametric bootstrap procedure to approximate these quantities.

Given $\mathbb{X}_n := (X'_1, \dots, X'_n)'$, let X^* and Y^* denote vectors of bootstrap random variables with values in \mathbb{R}^{d_1} and \mathbb{R}^{d_2} , respectively. To describe the dependence structure of the bootstrap sample, we define, in analogy to Definition 1.1,

$$\tau^*(Y^*, X^*, x_n) := \mathbb{E} \left(\sup_{f \in \mathcal{A}_1(\mathbb{R}^{d_1})} \left| \int_{\mathbb{R}^{d_1}} f(x) P_{X^*|Y^*}(dx) - \int_{\mathbb{R}^{d_1}} f(x) P_{X^*}(dx) \right| \middle| \mathbb{X}_n = x_n \right),$$

provided that $\mathbb{E}(\|X^*\| \mid \mathbb{X}_n = x_n) < \infty$. We make the following assumption concerning the bootstrap sample $(X_k^*)_k$ and the corresponding estimator $\mu_n^*(X_1^*, \dots, X_n^*)$ of the location parameter:

- (S3)(i) The sequence of bootstrap variables is stationary with probability tending to one. Additionally, $(X_{t_1}^{*'}, X_{t_2}^{*'})' \xrightarrow{d} (X_{t_1}', X_{t_2}')'$, $\forall t_1, t_2 \in \mathbb{N}$, holds true in probability.
- (ii) Conditionally on \mathbb{X}_n , the random variables $(X_k^*)_{k \in \mathbb{Z}}$ are τ -weakly dependent, i.e. there exist a sequence of coefficients $(\bar{\tau}_r)_{r \in \mathbb{N}}$ with $\sum_{r=1}^{\infty} r(\bar{\tau}_r)^{\delta^2} < \infty$ for some $\delta \in (0, 1)$ and a sequence of sets $(\mathfrak{X}_n)_{n \in \mathbb{N}}$ with $P(\mathbb{X}_n \in \mathfrak{X}_n) \rightarrow_{n \rightarrow \infty} 1$ such that the following property holds: For any sequence $(x_n)_{n \in \mathbb{N}}$ with $x_n \in \mathfrak{X}_n$, $\sup_{k \in \mathbb{N}} \mathbb{E}(\|X_k^*\| \mid \mathbb{X}_n = x_n) < \infty$ and

$$\tau_r^*(x_n) := \sup_{l \in \mathbb{N}} \frac{1}{l} \sup_{r+i \leq k_1, \dots, k_l} \{\tau^*((X_1^*, \dots, X_l^*), (X_{k_1}^{*'}, \dots, X_{k_l}^{*'})', x_n)\} \leq \bar{\tau}_r.$$

- (iii) $\mathbb{E}^*(\|X_1^*\|^{1+\varepsilon}) = O_p(1)$ for all $\varepsilon \in (0, 1/(3 - 2\delta))$.¹

- (iv) The sequence of estimators $\hat{\mu}_n^*$ admits the expansion

$$\hat{\mu}_n^* - \hat{\mu}_n = \frac{1}{n} \sum_{k=1}^n l(X_k^*, \hat{\mu}_n) + o_{p^*}(n^{-1/2}),$$

with $\mathbb{E}^* l(X_1^*, \hat{\mu}_n) = 0_d$ and $\mathbb{E}^* \|l(X_1^*, \hat{\mu}_n)\|^2 = O_p(1)$ for some $\nu > (2 - \delta)/(1 - \delta)$.²

Remark 2.3. (i) We do not assume the sequence of bootstrap variables to be stationary for all sufficiently large n . The reason for this is that for instance the ordinary ARCH(p)-bootstrap method with estimated parameters would violate this assumption; see also [29].

¹ Throughout the paper the expressions P^* and \mathbb{E}^* denote the bootstrap distribution and bootstrap expectation conditionally on X_1, \dots, X_n .

² $R_n^* = o_{p^*}(a_n)$ if $P^*(\|R_n^*\|/|a_n| > \varepsilon) \xrightarrow{P^*} 0$, $\forall \varepsilon > 0$.

- (ii) Neumann and Paparoditis [29] proved that in case of stationary Markov chains of finite order, the key for convergence of the finite-dimensional distributions is convergence of the conditional distributions, cf. their Lemma 4.2. In particular, they showed that parametric AR(p)- and ARCH(p)-bootstrap methods yield samples that satisfy (S2)(i). Also sieve bootstrap procedures for linear processes exhibit this property; see [9].

Intuitively, one might consider to use the bootstrap counterpart of \widehat{S}_n as the bootstrap test statistic. Unfortunately, it turns out that the bootstrap counterpart of the approximating V -statistic S_n is no longer degenerate in general. This is however not surprising since a similar problem occurs when Efron's bootstrap is applied to V -type statistics of i.i.d. data. Arcones and Giné [2] proposed to degenerate the kernel on the bootstrap side artificially to overcome this difficulty. We will proceed analogously, cf. the proof of Proposition 2.2 below. For this reason the additional term $-\Im[c^*(t)]$ has to be included in the integrand of the bootstrap statistic \widehat{S}_n^* . Here, c^* is the characteristic function of $X_1^* - \widehat{\mu}_n$, conditionally on \mathbb{X}_n . More precisely, we propose the following bootstrap algorithm:

- (1) Determine $\widehat{\mu}_n$ such that (S1)(iii) is satisfied.
- (2) Generate X_1^*, \dots, X_n^* such that (S3)(i)–(iii) hold.
- (3) Determine $\widehat{\mu}_n^*$ such that (S3)(iv) is satisfied.
- (4) Compute the bootstrap version of our test statistic:

$$\widehat{S}_n^* := \int_{\mathbb{R}^d} \left\{ \frac{1}{\sqrt{n}} \sum_{k=1}^n (\sin(t'(X_k^* - \widehat{\mu}_n^*)) - \Im[c^*(t)]) \right\}^2 w(t) dt.$$

- (5) Define the critical value t_α^* as the $(1 - \alpha)$ -quantile of the (conditional) distribution of \widehat{S}_n^* . Reject \mathcal{H}_0 if $\widehat{S}_n > t_\alpha^*$.
(In practice, steps (1)–(4) will be repeated B times, for some large B . The critical value will then be approximated by the $(1 - \alpha)$ -quantile of the empirical distribution associated with $\widehat{S}_{n,1}^*, \dots, \widehat{S}_{n,B}^*$.)

Employing results of bootstrap for V -statistics, we can verify consistency of the algorithm above. So far, bootstrap for degenerate V -statistics of dependent data has only been investigated by Leucht [25]. However, her findings cannot be applied directly in the present context but have to be slightly modified, cf. proof of Proposition 2.2.

Proposition 2.2. Suppose that assumptions (S1) and (S3) hold and that \widehat{S}_n^* , $n \in \mathbb{N}$, are generated via the aforementioned algorithm. Then, under \mathcal{H}_0 ,

$$\widehat{S}_n^* \xrightarrow{d} Z_1 \quad \text{in probability,}$$

as $n \rightarrow \infty$, where Z_1 is defined as in Theorem 2.1. Moreover, if $\text{var}(Z_1) > 0$,

$$\sup_{-\infty < x < \infty} |P^*(\widehat{S}_n^* \leq x) - P(\widehat{S}_n \leq x)| \xrightarrow{P} 0.$$

The latter assertion implies that, under the null hypothesis, the bootstrap-based test has asymptotically the correct size. Moreover, consistency and asymptotic unbiasedness under Pitman alternatives of the bootstrap-aided test can be verified. Although Proposition 2.1 merely yields unbiasedness, we conjecture that our test has non-trivial power against some alternatives of the structure $\mathcal{H}_{1,n}$. In the i.i.d. case for instance, one can verify non-trivial power for certain local alternatives based on the asymptotic results of U -statistics under contiguous alternatives of [19]. However, technical conditions were imposed that depend on the underlying distribution in a complicated way and that can hardly be checked here. Since things become even more involved in the case of dependent observations, we do not pursue these investigations here.

Corollary 2.1. Suppose that assumptions (S1) and (S3) are fulfilled and let $\alpha \in (0, 1)$. The proposed bootstrap test based on the algorithm above satisfies

$$\lim_{n \rightarrow \infty} P(\widehat{S}_n > t_\alpha^*) = \begin{cases} \alpha & \text{if } \mathcal{H}_0 \text{ is true and } \text{var}(Z_1) > 0, \\ 1 & \text{if } \mathcal{H}_1 \text{ is true.} \end{cases}$$

Under $\mathcal{H}_{1,n}$ and the additional assumptions of Proposition 2.1,

$$\liminf_{n \rightarrow \infty} P_{\mathcal{H}_{1,n}}(\widehat{S}_n > t_\alpha^*) \geq \alpha.$$

2.3. Numerical examples

Example 2.1. We illustrate the finite sample behaviour of our test by a simulation study. To this end, X_1, \dots, X_n ($n = 50, 70, 100$) were generated due to an AR(1) process with model equation

$$X_t = 0.6X_{t-1} + \epsilon_t$$

Table 1
Rejection frequencies.

n	$\varepsilon_k \sim \mathcal{N}(0, 1)$		$\varepsilon_k \sim \mathcal{L}og\mathcal{N}(0, 1)$		$\varepsilon_k \sim \chi_1^2$	
	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$
50	0.050	0.100	0.170	0.315	0.220	0.405
70	0.040	0.085	0.400	0.540	0.330	0.545
100	0.050	0.100	0.635	0.785	0.655	0.805

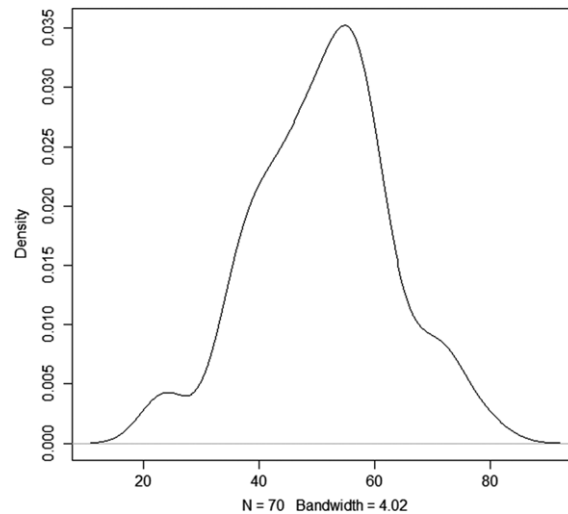


Fig. 1. Estimated density.

and i.i.d. innovations $(\varepsilon_k)_k$. We chose $\varepsilon_1 \sim \mathcal{N}(0, 1)$ in order to create a null scenario. For the investigation of the test under fixed alternatives, we drew the innovations from the log-normal distribution with parameters 0 and 1 and the χ_1^2 distribution. As a weight function for our test statistic we chose $w(t) = e^{-0.01t^2}$ which puts comparatively much mass around the origin, where the absolute value of the characteristic function is high. Another advantage of this choice is that the integral determining our test statistic can be easily calculated and therefore no numerical integration is required. The arithmetic mean was employed to estimate the potential centre of symmetry. We generated the bootstrap samples via classical residual-based AR(1) bootstrap. Of course, in this case the corresponding characteristic function c^* is unknown. We approximated this quantity by its empirical counterpart based on samples of size 1500 for $n = 50$, 1750 for $n = 70$ and 3000 for $n = 100$. The simulations were repeated 200 times, each with 200 bootstrap resamplings. All implementations were carried out with the aid of the statistical software package R; see [32]. Table 1 reports that our test keeps the desired size already in the case of moderate sample sizes. Still, in order to obtain good power properties larger sample sizes are required.

Example 2.2. As a real data example we consider Series F of length 70 in [8], i.e. yields from a batch chemical process, which fits an AR(2) process. Fig. 1 shows the corresponding density estimate generated by the function `density` of the software package R. It indicates skewness of the underlying distribution. In accordance, our test rejects the null hypothesis of symmetry. Here, we used the same weight function as in Example 2.1 and 500 bootstrap replications. While $\hat{S}_{70} = 29.61$, we obtain 16.98 for the bootstrap approximation of the upper 5% level.

3. A goodness-of-fit test for the marginal distribution of a time series

3.1. The test statistic and its asymptotics

We extend the test originally proposed by Fan [15] for i.i.d. random variables to weakly dependent observations. Suppose X_1, \dots, X_n are \mathbb{R}^d -valued observations with distribution P_X and consider the test problem

$$\mathcal{H}_0: P_X \in \{P^\theta \mid \theta \in \Theta\} \quad \text{vs.} \quad \mathcal{H}_1: P_X \notin \{P^\theta \mid \theta \in \Theta\},$$

$\Theta \subseteq \mathbb{R}^p$, which is equivalent to

$$\mathcal{H}_0: c_X = c(\cdot, \theta_0) \quad \text{for some } \theta_0 \in \Theta \quad \text{vs.} \quad \mathcal{H}_1: c_X \neq c(\cdot, \theta) \quad \forall \theta \in \Theta.$$

Here, c_X and $c(\cdot, \theta)$ denote the characteristic functions associated with P_X and P^θ , respectively. Fan [15] suggested a test statistic of the form

$$\widehat{G}_n = n \int_{\mathbb{R}^d} |c_n(t) - c(t, \widehat{\theta}_n)|^2 w(t) dt,$$

where $\widehat{\theta}_n$ is a consistent estimator of θ_0 and w is an appropriate weight function. In order to derive the limit distribution of \widehat{G}_n , we make the following assumptions:

(G1)(i) $(X_n)_n$ is a stationary, \mathbb{R}^d -valued, τ -dependent process with $\sum_{r=1}^{\infty} r(\tau_r)^{\delta^2} < \infty$ and $\mathbb{E}\|X_1\|^{(2-\delta)/(1-\delta)} < \infty$ for some $\delta \in (0, 1)$.

(ii) The sequence of estimators $\widehat{\theta}_n$ admits the expansion

$$\widehat{\theta}_n - \theta_0 = \frac{1}{n} \sum_{k=1}^n l(X_k, \theta_0) + o_P(n^{-1/2}), \quad (3.1)$$

where $\mathbb{E}l(X_1, \theta_0) = 0$ and $\mathbb{E}\|l(X_1, \theta_0)\|^{2\nu} < \infty$ for some $\nu > (2 - \delta)/(1 - \delta)$. Moreover, $\|l(x, \theta_0) - l(\bar{x}, \theta_0)\| \leq f_l(x, \bar{x}, \theta_0) \|x - \bar{x}\|$ where $f_l(\cdot, \cdot, \theta_0)$ is continuous and such that for some $A > 0$,

$$\sup_{Y_1, Y_2 \sim P_X} \mathbb{E} \max_{a \in [-A, A]^d} |f_l(Y_1 + a, Y_2, \theta_0)|^{2(2-\delta)/(1-\delta)} < \infty.$$

(iii) The weight function $w : \mathbb{R}^d \rightarrow \mathbb{R}$ is measurable, positive almost everywhere w.r.t. the Lebesgue measure on \mathbb{R}^d and satisfies $\int_{\mathbb{R}^d} (1 + \|t\|^2) w(t) dt < \infty$.

(iv) The function $c(t, \cdot)$ is twice continuously differentiable $\forall t \in \mathbb{R}^d$ with $\int_{\mathbb{R}^d} \|c^{(1)}(t, \theta_0)\|^{2\nu} w(t) dt < \infty$. Additionally, there exists a neighbourhood $U(\theta_0) \subseteq \Theta$ such that for all $\theta \in U(\theta_0)$ every element of $c^{(2)}(t, \theta)$ can be bounded by a symmetric function M with $\int_{\mathbb{R}^d} M^2(t) w(t) dt < \infty$.

The expansion (3.1) of $\widehat{\theta}_n - \theta_0$ is a standard assumption in the field of hypothesis testing. It is often satisfied when maximum-likelihood estimators or the method of moments are applied. Even though the smoothness assumption on the associated function l has a rather technical structure, it is satisfied e.g. by polynomial functions as long as the sample variables have sufficiently many finite moments. Under the null hypothesis, \widehat{G}_n can be approximated by a degenerate V -statistic and thus the limit distribution can be derived similarly to the previous section.

Theorem 3.1. Suppose that assumption (G1) holds. Then, under \mathcal{H}_0 ,

(i) $\widehat{G}_n - G_n = o_P(1)$, where

$$G_n = n \int_{\mathbb{R}^d} \left| c_n(t) - c(t, \theta_0) - \frac{1}{n} \sum_{j=1}^n [l(X_j, \theta_0)]' c^{(1)}(t, \theta_0) \right|^2 w(t) dt.$$

(ii) Moreover, there are families of constants $(\lambda_{j;k_1,k_2}^{(c)}(\theta_0))_{c,j,k_1,k_2}$, $(C_{j;k_1,k_2})_{j,k_1,k_2}$ and centred, jointly normal random variables $(Z_{j;k})_{j,k}$ such that

$$\begin{aligned} \widehat{G}_n \xrightarrow{d} Z_2 &:= \lim_{c \rightarrow \infty} \sum_{j=1}^{\infty} \sum_{k_1, k_2 \in \mathbb{Z}^d} \lambda_{j;k_1,k_2}^{(c)}(\theta_0) [Z_{j;k_1} Z_{j;k_2} + C_{j;k_1,k_2}] \\ &\quad + \mathbb{E} \int_{\mathbb{R}^d} |e^{it'X_1} - c(t, \theta_0) - [l(X_1, \theta_0)]' c^{(1)}(t, \theta_0)|^2 w(t) dt. \end{aligned}$$

Here, the r.h.s. converges in the L_2 -sense.

As in the previous section, the test statistic is asymptotically unbounded under fixed alternatives.

Lemma 3.1. Suppose that (G1)(i) and (iii) hold. Moreover, assume that there exists a constant $\theta_0 \in \Theta$ such that $\widehat{\theta}_n \xrightarrow{P} \theta_0$. Then, under \mathcal{H}_1 ,

$$P(\widehat{G}_n > K) \xrightarrow{n \rightarrow \infty} 1 \quad \forall K < \infty.$$

We conclude with a result concerning the behaviour of the test statistic under local alternatives

$$\begin{aligned} \mathcal{H}_{1,n}: \frac{dP_{X_{n,1}}}{dP_{X_1}} &= 1 + \frac{g_n}{\sqrt{n}}, \quad \text{where } \|g_n - g\|_{\infty} \xrightarrow{n \rightarrow \infty} 0 \text{ for some measurable, bounded function } g \text{ and with} \\ &\int_{\mathbb{R}^d} e^{it'x} g_n(x) P_{X_1}(dx) \neq 0 \text{ for some } t \in \mathbb{R}^d. \end{aligned}$$

Proposition 3.1. Suppose that the conditions (G1) and (S2) are satisfied. Additionally, assume that $\widehat{\theta}_{n,n} = \widehat{\theta}_n(X_{n,1}, \dots, X_{n,n}) = n^{-1} \sum_{k=1}^n l(X_{n,k}, \theta_0) + o_P(n^{-1/2})$ with

$$\sup_{n \in \mathbb{N}} \sup_{Y_{n,1}, Y_{n,2} \sim P_{X_{n,1}}} \mathbb{E} \max_{a \in [-A, A]^d} |f_l(Y_{n,1} + a, Y_{n,2})|^{2(2-\delta)/(1-\delta)} < \infty.$$

Then, under $\mathcal{H}_{1,n}$,

$$\begin{aligned} \widehat{G}_n \xrightarrow{d} Z_{2,loc} &:= \lim_{c \rightarrow \infty} \sum_{j=1}^{\infty} \sum_{k_1, k_2 \in \mathbb{Z}^d} \lambda_{j,k_1,k_2}^{(c)}(\theta_0) [\{Z_{j,k_1} + D_{j,k_1}\} \{Z_{j,k_2} + D_{j,k_2}\} + C_{j,k_1,k_2}] \\ &\quad + \mathbb{E} \int_{\mathbb{R}^d} |e^{it'X_1} - c(t, \theta_0) - [l(X_1, \theta_0)]' c^{(1)}(t, \theta_0)|^2 w(t) dt \end{aligned}$$

with a family of constants $(D_{j,k})_{j,k}$. If additionally $\text{var}(Z_2) > 0$, then

$$\liminf_{n \rightarrow \infty} [P_{\mathcal{H}_{1,n}}(\widehat{G}_n > x) - P_{\mathcal{H}_0}(\widehat{G}_n > x)] \geq 0 \quad \forall x \in \mathbb{R}.$$

3.2. Bootstrap validity

Again we propose to invoke a parametric bootstrap procedure to determine critical values of the test. For this purpose we assume:

(G2)(i) The sequence of bootstrap variables is stationary with probability tending to one. Additionally, (S3)(ii) holds and

$(X_{t_1}^*, X_{t_2}^*)' \xrightarrow{d} (Y_{t_1}', Y_{t_2}')', \forall t_1, t_2 \in \mathbb{N}$, in probability, where $(Y_n)_{n \in \mathbb{N}}$ satisfies (G1)(i), (ii) and $(Y_{t_1}', Y_{t_2}')' = (X_{t_1}', X_{t_2}')'$ under \mathcal{H}_0 .

(ii) $\mathbb{E}^* \|X_1^*\|^{(2-\delta)/(1-\delta)} \xrightarrow{P} \mathbb{E} \|X_1\|^{(2-\delta)/(1-\delta)}$.

(iii) The sequence of estimators $\widehat{\theta}_n^*$ admits the expansion

$$\widehat{\theta}_n^* - \widehat{\theta}_n = \frac{1}{n} \sum_{k=1}^n l(X_k^*, \widehat{\theta}_n) + o_{P^*}(n^{-1/2}),$$

where $\mathbb{E}^* l(X_1^*, \widehat{\theta}_n) = 0$ and $\mathbb{E}^* \|l(X_1^*, \widehat{\theta}_n)\|^{2\nu} = O_P(1)$ for some $\nu > (2-\delta)/(1-\delta)$. Additionally, $\|l(\bar{x}, \widehat{\theta}_n) - l(\bar{x}, \widehat{\theta}_n)\| \leq f_l(\bar{x}, \bar{x}, \widehat{\theta}_n) \|\bar{x} - \bar{x}\|$. Moreover, for some $A, K \in (0, \infty)$ there exists a sequence of sets $(\mathbb{X}_n)_n$ such that $P(\mathbb{X}_n \in \mathbb{X}_n) \rightarrow 1$ as $n \rightarrow \infty$ and $\forall (x_n)_n$ with $x_n \in \mathbb{X}_n$:

$$\mathbb{E}^* \left(\max_{a \in [-A, A]^d} |f_l(Y_1^* + a, Y_2^*, \widehat{\theta}_n)|^{2(2-\delta)/(1-\delta)} \mid \mathbb{X}_n = x_n \right) \leq K$$

for any Y_i^* with $Y_i^* \stackrel{d}{=} X_i^*$, $i = 1, 2$, conditionally on \mathbb{X}_n .

(iv) The first derivative of c w.r.t. θ satisfies $\int_{\mathbb{R}^d} \|c^{(1)}(t, \theta)\|^{2\nu} w(t) dt < \infty$, $\forall \theta \in U(\theta_0)$.

In the i.i.d. case, the validity of parametric bootstrap was derived by Fan [15]. She employed the bootstrap test statistic

$$\widetilde{G}_n^* = n \int_{\mathbb{R}^d} |c_n^*(t) - c(t, \widehat{\theta}_n^*)|^2 w(t) dt,$$

where c_n^* is the empirical counterpart of c_n , i.e. $c_n^*(t) = n^{-1} \sum_{k=1}^n e^{it'X_k^*}$. When bootstrapping the test statistic \widehat{G}_n under weak dependence, similar problems as in the previous section occur. Again the bootstrap counterpart \widetilde{G}_n^* of the approximating V -statistic G_n is not degenerate in general. In order to establish a consistent bootstrap method to determine critical values of the test statistic, we do therefore not use \widetilde{G}_n^* directly but include the additional term $c(t, \widehat{\theta}_n) - c^*(t)$ in its integrand. Here, c^* is the characteristic function of X_1^* , conditionally on X_1, \dots, X_n . We suggest the application of the following bootstrap algorithm:

- (1) Determine $\widehat{\theta}_n$ such that (G1)(ii) is satisfied.
- (2) Generate X_1^*, \dots, X_n^* such that (G2)(i) and (ii) hold.
- (3) Determine $\widehat{\theta}_n^*$ such that (G2)(iii) is satisfied.
- (4) Compute the bootstrap test statistic

$$\widehat{G}_n^* := n \int_{\mathbb{R}^d} |c_n^*(t) - c^*(t) + c(t, \widehat{\theta}_n) - c(t, \widehat{\theta}_n^*)|^2 w(t) dt.$$

- (5) Define the critical value t_α^* as the $(1-\alpha)$ -quantile of the (conditional) distribution of \widehat{G}_n^* . Reject \mathcal{H}_0 if $\widehat{G}_n > t_\alpha^*$.

Table 2
Rejection frequencies.

n	a	$\varepsilon_k \sim \mathcal{N}(0, (1 - a^2))$		$\varepsilon_k \sim \mathcal{L}og\mathcal{N}(0, 1)$		$\varepsilon_k \sim \chi_1^2$	
		$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$
100	0.6	0.050	0.090	0.460	0.640	0.460	0.575
200	0.6	0.045	0.085	0.925	0.950	0.825	0.895
100	−0.6	0.015	0.055	0.410	0.525	0.435	0.485
200	−0.6	0.030	0.080	0.830	0.880	0.675	0.740

Remark 3.1. The bootstrap characteristic function c^* is often unknown and has to be approximated by simulation. However, note that there are cases where the bootstrap variables can be generated such that $\hat{c}^*(t) = c(t, \hat{\theta}_n)$; see [35] and Section 3.3 below. In these situations \hat{G}_n^* coincides with the exact bootstrap counterpart \tilde{G}_n^* of \hat{G}_n .

Proposition 3.2. Suppose that the assumptions (G1) and (G2) are satisfied. Then, under \mathcal{H}_0 ,

$$\hat{G}_n^* \xrightarrow{d} Z_2 \quad \text{in probability,}$$

as $n \rightarrow \infty$, where Z_2 is defined as in Theorem 3.1. Moreover, if $\text{var}(Z_2) > 0$,

$$\sup_{-\infty < x < \infty} |P^*(\hat{G}_n^* \leq x) - P(\hat{G}_n \leq x)| \xrightarrow{P} 0.$$

The proposition assures that the bootstrap-based test is asymptotically of correct size under \mathcal{H}_0 . Additionally, it is consistent w.r.t. fixed alternatives and asymptotically unbiased against Pitman local alternatives. The asymptotic behaviour of our bootstrap-aided test can be summarized as follows.

Corollary 3.1. Suppose that the assumptions (G1) and (G2) are fulfilled and let $\alpha \in (0, 1)$. The proposed bootstrap test based on the algorithm above satisfies

$$\lim_{n \rightarrow \infty} P(\hat{G}_n > t_\alpha^*) = \begin{cases} \alpha & \text{if } \mathcal{H}_0 \text{ is true and } \text{var}(Z_2) > 0, \\ 1 & \text{if } \mathcal{H}_1 \text{ is true.} \end{cases}$$

Moreover, under the additional assumptions of Proposition 3.1,

$$\liminf_{n \rightarrow \infty} P_{\mathcal{H}_{1,n}}(\hat{G}_n > t_\alpha^*) \geq \alpha.$$

3.3. Numerical examples

Example 3.1. Since the effect of artificial degeneration of the bootstrap statistic has been already illuminated in case of the test for symmetry, we illustrate the finite sample behaviour of our goodness-of-fit test in a situation now, where this is not required. Suppose that we have observations X_1, \dots, X_n from a stationary AR(1) process, that is,

$$X_k = aX_{k-1} + \epsilon_k, \quad \text{for some (unknown) } a \in (-1, 1)$$

with i.i.d. innovations $(\epsilon_k)_k$. Let us consider the problem

$$\begin{aligned} \mathcal{H}_0: X_1 &\sim \mathcal{N}(\mu, \sigma^2) \quad \text{for some } \mu \in \mathbb{R}, \sigma^2 > 0 \quad \text{vs.} \\ \mathcal{H}_1: X_1 &\not\sim \mathcal{N}(\mu, \sigma^2) \quad \forall \mu \in \mathbb{R}, \sigma^2 > 0. \end{aligned}$$

In order to provide null scenarios, we generated samples X_1, \dots, X_n of size $n = 100$ and $n = 200$ from AR(1) processes with $a = 0.6$ and $a = -0.6$ and $\epsilon_1 \sim \mathcal{N}(0, 1 - a^2)$. Thus $X_1 \sim \mathcal{N}(0, 1)$. We drew the innovations from the log-normal distribution with parameters 0 and 1 and the χ_1^2 distribution to discuss the behaviour of the test under fixed alternatives. On the bootstrap side we generated an AR(1) process with estimated parameter $\hat{a}(X_1, \dots, X_n)$ and innovations that are distributed according to $\mathcal{N}(\hat{\mu}(1 - \hat{a}), \hat{\sigma}^2(1 - \hat{a}^2))$, where \hat{a} denotes the Yule–Walker estimator of a , $\hat{\mu}$ is the arithmetic mean and $\hat{\sigma}^2$ the empirical variance of the underlying sample. This approach guarantees that c^* and $c(\cdot, (\hat{\mu}, \hat{\sigma}^2))$ coincide. This in turn implies that \hat{G}_n^* simplifies to \tilde{G}_n^* . Here, the weight function was chosen as $w(t) = e^{-0.0005t^2}$, which puts sufficient weight on the region where the absolute value of the characteristic function of the underlying sample is significantly greater than zero. We replicated the simulations 200 times each with 200 bootstrap resamplings. Table 2 shows that our procedure is a little conservative for moderate samples. Computational requirements prevented us from considering larger sample sizes. However, we still obtained very satisfactory results. In particular the power rises substantially with increasing sample size.

Example 3.2. We apply our test to the data that we already used in [Example 2.2](#). The test for normality in $AR(p)$ processes of [22] does not reject the hypothesis of normality. In contrast our test (with 500 bootstrap replications) rejects this null hypothesis. Using the same weight function as in [Example 3.1](#), we obtain $\hat{G}_{70} = 132.90$ while the upper 5%-level of the bootstrap distribution is equal to 50.05.

4. Proofs

Throughout this section, C denotes a generic finite constant that may change its value even within a single calculation.

4.1. Proofs of the main results

Proof of Lemma 2.1. Define

$$\tilde{S}_n := \int_{\mathbb{R}^d} \left(\frac{1}{\sqrt{n}} \sum_{k=1}^n [\sin(t'(X_k - \mu_0)) - (\hat{\mu}_n - \mu_0)' t \cos(t'(X_k - \mu_0))] \right)^2 w(t) dt.$$

Step 1. $\hat{S}_n - \tilde{S}_n = o_p(1)$.

The application of an addition formula for trigonometric functions yields

$$\begin{aligned} \hat{S}_n - \tilde{S}_n &= \int_{\mathbb{R}^d} \frac{1}{n} \sum_{j,k=1}^n \sin(t'(X_j - \mu_0)) \sin(t'(X_k - \mu_0)) [\cos^2(t'(\mu_0 - \hat{\mu}_n)) - 1] w(t) dt \\ &\quad + \int_{\mathbb{R}^d} \frac{2}{n} \sum_{j,k=1}^n \{\sin(t'(X_j - \mu_0)) \cos(t'(X_k - \mu_0)) \\ &\quad \times [\sin(t'(\mu_0 - \hat{\mu}_n)) \cos(t'(\mu_0 - \hat{\mu}_n)) + t'(\hat{\mu}_n - \mu_0)]\} w(t) dt \\ &\quad + \int_{\mathbb{R}^d} \frac{1}{n} \sum_{j,k=1}^n \{\cos(t'(X_j - \mu_0)) \cos(t'(X_k - \mu_0)) [\sin^2(t'(\mu_0 - \hat{\mu}_n)) - (t'(\hat{\mu}_n - \mu_0))^2]\} w(t) dt \\ &=: T_1 + T_2 + T_3. \end{aligned}$$

Since $\sqrt{n}(\hat{\mu}_n - \mu_0) = n^{-1/2} \sum_{k=1}^n l(X_k, \mu_0) + o_p(1)$, [Lemma 4.1](#) implies $\sqrt{n}(\hat{\mu}_n - \mu_0) = O_p(1)$. This leads to $|\cos^2(t'(\mu_0 - \hat{\mu}_n)) - 1| \leq \|t\| o_p(1)$, which in turn implies that $T_1 \leq o_p(1) \int_{\mathbb{R}^d} [n^{-1/2} \sum_{k=1}^n \sin(t'(X_k - \mu_0))]^2 \|t\| w(t) dt$. In order to prove that the integral is of order $O_p(1)$, it remains to show that for some $C < \infty$ the inequality

$$\int_{\mathbb{R}^d} \mathbb{E} \left[\frac{1}{\sqrt{n}(1 + \|t\|)} \sum_{k=1}^n \sin(t'(X_k - \mu_0)) \right]^2 \|t\| (1 + \|t\|) w(t) dt < C, \quad \forall n \in \mathbb{N},$$

holds. Applying [Lemma 4.1](#) with $g_n(x) = \sin(t'(x - \mu_0))/(1 + \|t\|)$, we obtain $\sup_t \mathbb{E}[(n(1 + \|t\|))^{-1/2} \sum_{k=1}^n \sin(t'(X_k - \mu_0))]^2 < \infty$. The integrability constraint on the function w finally leads to $T_1 = o_p(1)$.

Furthermore, the equality $2 \sin(t'(\mu_0 - \hat{\mu}_n)) \cos(t'(\mu_0 - \hat{\mu}_n)) = \sin(2t'(\mu_0 - \hat{\mu}_n))$ and a Taylor expansion of $\sin(2t'(\mu_0 - \hat{\mu}_n))$ in the origin yield

$$\begin{aligned} |T_2| &\leq \int_{\mathbb{R}^d} \left| \sum_{k=1}^n \sin(t'(X_k - \mu_0)) \right| |\sin(2t'(\mu_0 - \hat{\mu}_n)) - 2t'(\mu_0 - \hat{\mu}_n)| w(t) dt \\ &\leq 2 \int_{\mathbb{R}^d} \left| \sum_{k=1}^n \sin(t'(X_k - \mu_0)) \right| [t'(\mu_0 - \hat{\mu}_n)]^2 w(t) dt. \end{aligned}$$

In the same manner as before, we get $|T_2| = o_p(1)$.

Finally, T_3 can be approximated by applying the identity $2 \sin^2(t'(\mu_0 - \hat{\mu}_n)) = 1 - \cos(2t'(\mu_0 - \hat{\mu}_n))$ and by Taylor expansion of $\cos(2t'(\mu_0 - \hat{\mu}_n))$ in the origin:

$$|T_3| \leq n \int_{\mathbb{R}^d} \left| \frac{1}{2} [1 - \cos(2t'(\mu_0 - \hat{\mu}_n))] - [t'(\mu_0 - \hat{\mu}_n)]^2 \right| w(t) dt \leq O_p(n^{-1/2}).$$

Step 2. $\tilde{S}_n - S_n = o_p(1)$.

We split up

$$\begin{aligned}\tilde{S}_n - S_n &= \int_{\mathbb{R}^d} \frac{2}{n} \sum_{j,k=1}^n \sin(t'(X_j - \mu_0)) [c_{X-\mu_0}(t) l(X_k, \mu_0)' t - (\hat{\mu}_n - \mu_0)' t \cos(t(X_k - \mu_0))] w(t) dt \\ &\quad + \int_{\mathbb{R}^d} \frac{1}{n} \sum_{j,k=1}^n [\cos(t'(X_j - \mu_0)) \cos(t'(X_k - \mu_0)) [t'(\hat{\mu}_n - \mu_0)]^2 - c_{X-\mu_0}^2(t) l(X_j, \mu_0)' t l(X_k, \mu_0)' t] w(t) dt \\ &= R_1 + R_2.\end{aligned}$$

Concerning the first summand one gets

$$\begin{aligned}|R_1| &\leq \int_{\mathbb{R}^d} \left| \frac{2}{\sqrt{n}} \sum_{j=1}^n \sin(t'(X_j - \mu_0)) \right| |c_{X-\mu_0}(t)| \left| \frac{1}{\sqrt{n}} \sum_{k=1}^n l(X_k, \mu_0)' t - \sqrt{n}(\hat{\mu}_n - \mu_0)' t \right| w(t) dt \\ &\quad + \int_{\mathbb{R}^d} \left| \frac{2}{\sqrt{n}} \sum_{j=1}^n \sin(t'(X_j - \mu_0)) \right| \left| c_{X-\mu_0}(t) - \frac{1}{n} \sum_{k=1}^n \cos(t'(X_k - \mu_0)) \right| |\sqrt{n}(\hat{\mu}_n - \mu_0)' t| w(t) dt \\ &= o_P(1)\end{aligned}$$

by (S1)(iii) and Lemma 4.1. The latter result additionally implies

$$\begin{aligned}|R_2| &= o_P(1) + \left| \int_{\mathbb{R}^d} \frac{1}{n} \sum_{i,j=1}^n l(X_i, \mu_0)' t l(X_j, \mu_0)' t \left[\frac{1}{n^2} \sum_{k,m=1}^n \cos(t'(X_k - \mu_0)) \cos(t'(X_m - \mu_0)) - c_{X-\mu_0}^2(t) \right] w(t) dt \right| \\ &= o_P(1) + O_P(1) \int_{\mathbb{R}^d} \left| \frac{1}{n} \sum_{k=1}^n \cos(t'(X_k - \mu_0)) - c_{X-\mu_0}(t) \right| \|t\|^2 w(t) dt,\end{aligned}$$

which is asymptotically negligible. \square

Proof of Theorem 2.1. In view of Lemma 2.1 it suffices to show that $S_n \xrightarrow{d} Z_1$. The statistic S_n is of degenerate V -type with symmetric kernel $h = h_{\mu_0}$ given by

$$h_{\mu_0}(x, y) = \int_{\mathbb{R}^d} [\sin(t'(x - \mu_0)) - c_{X-\mu_0}(t) t' l(x, \mu_0)] [\sin(t'(y - \mu_0)) - c_{X-\mu_0}(t) t' l(y, \mu_0)] w(t) dt.$$

We obtain its limit from Theorem 2.2 in [25] if her assumptions (A1), (A2) and (A4) are satisfied. The first condition follows from (S1)(i). Obviously, h_{μ_0} is degenerate under the null hypothesis, i.e. $\mathbb{E} h_{\mu_0}(x, X_1) = 0 \forall x \in \mathbb{R}^d$, and satisfies $\sup_{k \in \mathbb{N}} \mathbb{E} |h_{\mu_0}(X_1, X_{1+k})|^v + \mathbb{E} |h_{\mu_0}(X_1, \tilde{X}_1)|^v < \infty$. This implies (A2), where \tilde{X}_1 is an i.i.d. copy of X_1 . Moreover, h_{μ_0} has the following continuity property:

$$\begin{aligned}|h_{\mu_0}(x, y) - h_{\mu_0}(\bar{x}, \bar{y})| &\leq C(1 + \|l(x, \mu_0)\| + \|l(y, \mu_0)\| + \|l(\bar{x}, \mu_0)\| + \|l(\bar{y}, \mu_0)\|)(\|x - \bar{x}\| + \|y - \bar{y}\|) \\ &=: f(x, \bar{x}, y, \bar{y})(\|x - \bar{x}\| + \|y - \bar{y}\|).\end{aligned}$$

Due to stationarity, Lipschitz continuity of l and the moment constraints on l and X_1 , we obtain by Hölder's inequality for $\eta := 1/(1 - \delta)$ and some $a > 0$,

$$\begin{aligned}\sup_{Y_1, \dots, Y_5 \sim P_X} \mathbb{E} \left(\max_{c_1, c_2 \in [-a, a]^d} f(Y_1, Y_2 + c_1, Y_3, Y_4 + c_2)^\eta \|Y_5\| \right) \\ \leq C[1 + (\mathbb{E} \|l(X_1, \mu_0)\|^{2v})^{\eta/(2v)} (\mathbb{E} \|X_1\|^{2v/(2v-\eta)})^{(2v-\eta)/(2v)}] < \infty.\end{aligned}$$

To this end, also note that $2v/(2v - \eta) < 1 + (3 - 2\delta)^{-1}$ which actually implies $\mathbb{E} \|X_1\|^{2v/(2v-\eta)} < \infty$ by virtue of (S1)(ii). Consequently, (A4) of [25] holds and her Theorem 2.2 can be applied in order to determine the asymptotic distribution of S_n : Let ϕ and ψ denote \mathbb{R} -valued, Lipschitz continuous, compactly supported scale and wavelet functions associated with a one-dimensional multiresolution analysis and such that $\int_{-\infty}^{\infty} \phi(x) dx = 1$ and $\int_{-\infty}^{\infty} \psi(x) dx = 0$. Set

$$\varphi^{(i)} := \begin{cases} \phi & \text{for } i = 0, \\ \psi & \text{for } i = 1 \end{cases}$$

and $\Phi_{j,k}^{(e)}(x) = 2^{jd/2} \prod_{i=1}^d \varphi^{(e_i)}(2^j x_i - k_i)$ for $x \in \mathbb{R}^d$, $k \in \mathbb{Z}^d$, $e \in \{0, 1\}^d$. Let $h_{\mu_0}^{(c)}$ denote the truncated version of h_{μ_0} defined as the degenerate counterpart of $\tilde{h}_{\mu_0}^{(c)}$,

$$\tilde{h}_{\mu_0}^{(c)} := \begin{cases} h_{\mu_0}(x, y) & \text{for } |h_{\mu_0}(x, y)| \leq c_h, \\ -c_h & \text{for } h_{\mu_0}(x, y) < -c_h, \\ c_h & \text{for } h_{\mu_0}(x, y) > c_h \end{cases}$$

with $c_h := \max_{x,y \in [-c,c]^d} |h_{\mu_0}(x,y)|$. The limit distributions of [25] and the present theorem coincide when we define

$$\gamma_{j;k_1+e_1,k_2+e_2}^{(c)}(\mu_0) := \begin{cases} \iint_{\mathbb{R}^d \times \mathbb{R}^d} h_{\mu_0}^{(c)}(x,y) \Phi_{j;k_1/2}^{(e_1)}(x) \Phi_{j;k_2/2}^{(e_2)}(y) dx dy & \text{for } \frac{k_1}{2}, \frac{k_2}{2} \in \mathbb{Z}^d, (e'_1, e'_2)' \in \{0,1\}^{2d} \setminus \{0_{2d}\} \\ 0 & \text{or } \frac{k_1}{2}, \frac{k_2}{2} \in \mathbb{Z}^d, (e'_1, e'_2)' = 0_{2d}, j=0, \\ & \text{else.} \end{cases} \quad (4.1)$$

$C_{j;k_1+e_1,k_2+e_2} := \text{cov}(\Phi_{j;k_1/2}^{(e_1)}(X_1), \Phi_{j;k_2/2}^{(e_2)}(X_1))$ and if the family $(Z_{j;k})_{j,k}$ of jointly normal random variables is chosen such that

$$\begin{aligned} \text{cov}(Z_{j_1;k_1+e_1}, Z_{j_2;k_2+e_2}) &= \text{cov}(\Phi_{j_1;k_1/2}^{(e_1)}(X_1), \Phi_{j_2;k_2/2}^{(e_2)}(X_1)) \\ &+ \sum_{t=2}^{\infty} [\text{cov}(\Phi_{j_1;k_1/2}^{(e_1)}(X_1), \Phi_{j_2;k_2/2}^{(e_2)}(X_t)) + \text{cov}(\Phi_{j_1;k_1/2}^{(e_1)}(X_t), \Phi_{j_2;k_2/2}^{(e_2)}(X_1))]. \quad \square \end{aligned}$$

Proof of Lemma 2.2. Under \mathcal{H}_1 and the assumptions concerning w , $\int_{\mathbb{R}^d} [\Im(c_{X-\mu_0}(t))]^2 w(t) dt > 0$ holds true. Thus, it is sufficient to show that $P(|n^{-1}\widehat{S}_n - \int_{\mathbb{R}^d} [\Im(c_{X-\mu_0}(t))]^2 w(t) dt| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0 \forall \varepsilon > 0$ in order to verify the claim. According to the integrability assumption on w , it remains to prove that

$$P\left(\int_{[-T,T]^d} |\Im(e^{it'\widehat{\mu}_n} c_n(t)) - \Im(c_{X-\mu_0}(t))| w(t) dt > \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0 \quad \forall \varepsilon > 0. \quad (4.2)$$

Clearly, $\int_{[-T,T]^d} (\cdot)(t) w(t) dt$ is a continuous mapping from $C([-T,T]^d)$ to \mathbb{R}^+ , both endowed with the corresponding uniform metrics. The continuous mapping theorem implies (4.2) if $(n^{-1} \sum_{j=1}^n \sin[t'(X_j - \widehat{\mu}_n)])_{t \in [-T,T]^d} \xrightarrow{P} \Im(c_{X-\mu_0}(t))_{t \in [-T,T]^d}$ holds true. To this end, we first show pointwise convergence $n^{-1} \sum_{j=1}^n \sin[t'(X_j - \widehat{\mu}_n)] \xrightarrow{P} \Im(c_{X-\mu_0}(t))$ for any fixed $t \in \mathbb{R}^d$. Afterwards it remains to prove that

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P\left(\sup_{\|s-t\| < \delta} \left| \frac{1}{n} \sum_{j=1}^n \{\sin[s'(X_j - \widehat{\mu}_n)] - \Im(c_{X-\mu_0}(s)) - \sin[t'(X_j - \widehat{\mu}_n)] + \Im(c_{X-\mu_0}(t))\} \right| \geq \varepsilon\right) = 0 \quad (4.3)$$

for all $\varepsilon > 0$. In order to derive pointwise convergence, note that

$$\left| \frac{1}{n} \sum_{j=1}^n \sin[t'(X_j - \widehat{\mu}_n)] - \Im(c_{X-\mu_0}(t)) \right| \leq \left| \frac{1}{n} \sum_{j=1}^n \sin[t'(X_j - \mu_0)] - \Im(c_{X-\mu_0}(t)) \right| + o_P(1).$$

The remaining sum converges to zero in probability due to the law of large numbers; see Lemma 5.1 of [25]. For proving the relation (4.3), it suffices to verify that

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P\left(\sup_{\|s-t\| < \delta} \left| \frac{1}{n} \sum_{j=1}^n \{\sin[s'(X_j - \widehat{\mu}_n)] - \sin[t'(X_j - \widehat{\mu}_n)]\} \right| \geq \frac{\varepsilon}{2}\right)$$

vanishes according to the uniform continuity of the characteristic function. This term can be bounded by

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P\left(\delta \|\widehat{\mu}_n - \mu_0\| \geq \frac{\varepsilon}{4}\right) + \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P\left(\frac{\delta}{n} \sum_{j=1}^n \|X_j - \mu_0\| \geq \frac{\varepsilon}{4}\right) = 0.$$

The latter probability vanishes asymptotically since $n^{-1} \sum_{j=1}^n \|X_j - \mu_0\| \xrightarrow{P} \mathbb{E}\|X_1 - \mu_0\|$ under (S1)(i). \square

Proof of Proposition 2.1. This is an immediate consequence of Proposition 4.1 and Lemma 4.3. \square

Proof of Proposition 2.2. Proceeding as in the proof of Lemma 2.1, we obtain

$$\begin{aligned} \widehat{S}_n^* &= \int_{\mathbb{R}^d} \left(\frac{1}{\sqrt{n}} \sum_{k=1}^n \sin(t'(X_k^* - \widehat{\mu}_n)) - \Im[c^*(t)] - (\widehat{\mu}_n^* - \widehat{\mu}_n)' t \cos(t'(X_k^* - \widehat{\mu}_n)) \right)^2 w(t) dt + o_{P^*}(1) \\ &= \int_{\mathbb{R}^d} \left(\frac{1}{\sqrt{n}} \sum_{k=1}^n \sin(t'(X_k^* - \widehat{\mu}_n)) - \Im[c^*(t)] - c_{X-\mu_0}(t) t' l(X_k^*, \widehat{\mu}_n) \right)^2 w(t) dt + o_{P^*}(1). \end{aligned}$$

Therefore, it suffices to show that $S_n^* \xrightarrow{d} Z_1$ in probability, where

$$S_n^* := \int_{\mathbb{R}^d} \left(\frac{1}{\sqrt{n}} \sum_{k=1}^n \sin(t'(X_k^* - \hat{\mu}_n)) - \Im[c^*(t)] - c_{X-\mu_0}(t) t' l(X_k^*, \hat{\mu}_n) \right)^2 w(t) dt.$$

The direct application of Theorem 3.1 of [25] is not possible since S_n^* is not the bootstrap counterpart of S_n , i.e. the kernel functions of both statistics do not coincide. However, her proof of this result implies that $n\bar{V}_n^* \xrightarrow{d} Z_1$, in probability, where

$$\begin{aligned} \bar{V}_n^* := & \frac{1}{n^2} \sum_{j,k=1}^n \left[h(X_j^*, X_k^*, \hat{\mu}_n) - \int_{\mathbb{R}^d} h(x, X_k^*, \hat{\mu}_n) P_{X_1^*|X_1, \dots, X_n}(dx) - \int_{\mathbb{R}^d} h(X_j^*, y, \hat{\mu}_n) P_{X_1^*|X_1, \dots, X_n}(dy) \right. \\ & \left. + \iint_{\mathbb{R}^d \times \mathbb{R}^d} h(x, y, \hat{\mu}_n) P_{X_1^*|X_1, \dots, X_n}(dx) P_{X_1^*|X_1, \dots, X_n}(dy) \right] \end{aligned}$$

and

$$h(x, y, \hat{\mu}_n) = \int_{\mathbb{R}^d} [\sin(t'(x - \hat{\mu}_n)) - c_{X-\mu_0}(t) t' l(x, \hat{\mu}_n)] [\sin(t'(y - \hat{\mu}_n)) - c_{X-\mu_0}(t) t' l(y, \hat{\mu}_n)] w(t) dt.$$

Actually, $n\bar{V}_n^*$ and S_n^* coincide which finally yields the first assertion of the proposition. The second one is a consequence of the previous part if Z_1 has a continuous distribution function. This in turn follows from $\text{var}(Z_1) > 0$ by Lemma 3.2 of [25]. \square

Proof of Corollary 2.1. Step 1. $\lim_{n \rightarrow \infty} P(\hat{S}_n > t_\alpha^*) = \alpha$.

This is a consequence of Proposition 2.2.

Step 2. $\lim_{n \rightarrow \infty} P(\hat{S}_n > t_\alpha^*) = 1$.

Note that $\hat{S}_n^* - \bar{S}_n^* = o_{P^*}(1)$ holds under \mathcal{H}_1 . Here, the statistic \bar{S}_n^* is the bootstrap counterpart of a degenerate V-statistic \bar{S}_n with kernel

$$\begin{aligned} \bar{h}_{\mu_0}(x, y) := & \int_{\mathbb{R}^d} ([\sin(t'(x - \mu_0)) - \Im[c_{X-\mu_0}(t)] - \Re[c_{X-\mu_0}(t)] t' l(x, \mu_0)] \\ & \times [\sin(t'(y - \mu_0)) - \Im[c_{X-\mu_0}(t)] - \Re[c_{X-\mu_0}(t)] t' l(y, \mu_0)]) w(t) dt. \end{aligned}$$

Thus, the asymptotic $(1 - \alpha)$ -quantiles, $\alpha \in (0, 1)$, of \bar{S}_n are bounded. One obtains equivalence of the limits of \bar{S}_n and \bar{S}_n^* in analogy to Proposition 2.2. Consequently, the bootstrap quantiles are bounded in probability and the claim follows immediately from Lemma 2.2.

Step 3. $\liminf_{n \rightarrow \infty} P_{\mathcal{H}_{1,n}}(\hat{S}_n > t_\alpha^*) \geq \alpha$.

First note that the bootstrap algorithm imitates the null situation. In conjunction with Proposition 2.2 and the continuity of the distribution function of Z_1 , we obtain

$$P(t_\alpha^* \in [t_{\alpha+\delta}, t_{\alpha-\delta}]) \xrightarrow{n \rightarrow \infty} 1 \quad \forall \delta > 0 \text{ with } \alpha \pm \delta \in (0, 1), \quad (4.4)$$

where t_x denotes the $(1 - x)$ -quantile of the distribution of Z_1 . Let $\varepsilon > 0$ arbitrary but fixed, then we have to show $\liminf_{n \rightarrow \infty} P_{\mathcal{H}_{1,n}}(\hat{S}_n > t_\alpha^*) - \alpha \geq -\varepsilon$. According to step 1 of the proof it suffices to show that $\liminf_{n \rightarrow \infty} [P_{\mathcal{H}_{1,n}}(\hat{S}_n > t_\alpha^*) - P_{\mathcal{H}_0}(\hat{S}_n > t_\alpha^*)] > -\varepsilon$. The application of (4.4) and Proposition 2.1 lead to

$$\begin{aligned} \liminf_{n \rightarrow \infty} [P_{\mathcal{H}_{1,n}}(\hat{S}_n > t_\alpha^*) - P_{\mathcal{H}_0}(\hat{S}_n > t_\alpha^*)] & \geq \lim_{n \rightarrow \infty} [P_{\mathcal{H}_0}(\hat{S}_n > t_{\alpha-\delta}) - P_{\mathcal{H}_0}(\hat{S}_n > t_{\alpha+\delta})] \\ & = P(Z_1 > t_{\alpha-\delta}) - P(Z_1 > t_{\alpha+\delta}) \end{aligned}$$

which is greater than $-\varepsilon$ for any sufficiently small $\delta > 0$ as the cumulative distribution function of Z_1 is continuous. \square

Proof of Theorem 3.1. (i) Define

$$\tilde{G}_n := n \int_{\mathbb{R}^d} |c_n(t) - c(t, \theta_0) - (\hat{\theta}_n - \theta_0)' c^{(1)}(t, \theta_0)|^2 w(t) dt.$$

Step 1. $\hat{G}_n = \tilde{G}_n + o_P(1)$.

Recall that for any characteristic function c the equality $\overline{c(t)} = c(-t)$ holds for all $t \in \mathbb{R}^d$, where the bar denotes the complex conjugate. Taylor expansion gives

$$\begin{aligned} \hat{G}_n - \tilde{G}_n = & -\frac{n}{2} \int_{\mathbb{R}^d} [c_n(t) - c(t, \theta_0)] (\hat{\theta}_n - \theta_0)' c^{(2)}(-t, \tilde{\theta}) (\hat{\theta}_n - \theta_0) w(t) dt \\ & - \frac{n}{2} \int_{\mathbb{R}^d} [c_n(-t) - c(-t, \theta_0)] (\hat{\theta}_n - \theta_0)' c^{(2)}(t, \tilde{\theta}) (\hat{\theta}_n - \theta_0) w(t) dt \end{aligned}$$

$$\begin{aligned}
& + \frac{n}{2} \int_{\mathbb{R}^d} (\widehat{\theta}_n - \theta_0)' c^{(1)}(t, \theta_0) (\widehat{\theta}_n - \theta_0)' c^{(2)}(-t, \widetilde{\theta}) (\widehat{\theta}_n - \theta_0) w(t) dt \\
& + \frac{n}{2} \int_{\mathbb{R}^d} (\widehat{\theta}_n - \theta_0)' c^{(1)}(-t, \theta_0) (\widehat{\theta}_n - \theta_0)' c^{(2)}(t, \widetilde{\theta}) (\widehat{\theta}_n - \theta_0) w(t) dt \\
& + \frac{n}{4} \int_{\mathbb{R}^d} |(\widehat{\theta}_n - \theta_0)' c^{(2)}(t, \widetilde{\theta}) (\widehat{\theta}_n - \theta_0)|^2 w(t) dt
\end{aligned}$$

with some random $\widetilde{\theta}$ between $\widehat{\theta}_n$ and θ_0 . Since $\widehat{\theta}_n$ is a \sqrt{n} -consistent estimator, the latter three summands vanish asymptotically due to assumption (G1). The analyses of the remaining expressions are equal to each other. The absolute value of the first term can be bounded from above by

$$o_P(1) + O_P(1) \left(\int_{\mathbb{R}^d} |c_n(t) - c(t, \theta_0)|^2 w(t) dt \right)^{1/2} \left(\int_{\mathbb{R}^d} M^2(-t) w(t) dt \right)^{1/2}.$$

Similar to the consideration of the quantity T_1 in the proof of Lemma 2.1, asymptotic negligibility of the latter expression is a consequence of

$$\begin{aligned}
\mathbb{E}|c_n(t) - c(t, \theta_0)|^2 & \leq \frac{2}{n} + \frac{2}{n^2} \sum_{j < k} \mathbb{E}[\sin(t' X_j) - \Im c(t, \theta_0)][\sin(t' X_k) - \sin(t' \widetilde{X}_k)] \\
& \quad + \frac{2}{n^2} \sum_{j < k} \mathbb{E}[\cos(t' X_j) - \Re c(t, \theta_0)][\cos(t' X_k) - \cos(t' \widetilde{X}_k)] \\
& \leq \frac{2}{n} + \frac{C\|t\|}{n} \sum_{r=1}^{\infty} \tau_r.
\end{aligned} \tag{4.5}$$

Here, \widetilde{X}_k denotes a suitable copy of X_k that is independent of X_j .

Step 2. $\widehat{G}_n = G_n + o_P(1)$.

This follows easily from (4.5) in conjunction with (G1)(ii) and (iv).

- (ii) Under the assumption (G1) this assertion is an immediate consequence of part (i) in conjunction with Theorem 2.2 of [25]. We just remark that G_n can be formulated as a degenerate V-statistic with kernel

$$h_{\theta_0}(x, y) := \int_{\mathbb{R}^d} [\Re[g(x, t, \theta_0)] \Re[g(y, t, \theta_0)] + \Im[g(x, t, \theta_0)] \Im[g(y, t, \theta_0)] w(t) dt \tag{4.6}$$

and $g(x, t, \theta_0) = e^{it'x} - c(t, \theta_0) - [l(x, \theta_0)]' c^{(1)}(t, \theta_0)$. The function h_{θ_0} satisfies the moment constraint

$$\sup_{k \in \mathbb{N}} \mathbb{E}|h_{\theta_0}(X_1, X_{1+k})|^v + \mathbb{E}|h_{\theta_0}(X_1, \widetilde{X}_1)|^v < \infty,$$

where \widetilde{X}_1 is an independent copy of X_1 . Furthermore, note that h exhibits the continuity property

$$\begin{aligned}
|h_{\theta_0}(x, y) - h_{\theta_0}(\bar{x}, \bar{y})| & \leq C(\theta_0) \{ (1 + f_l(x, \bar{x}, \theta_0))(1 + \|l(y, \theta_0)\| + \|l(\bar{y}, \theta_0)\|) \\
& \quad + (1 + f_l(y, \bar{y}, \theta_0))(1 + \|l(x, \theta_0)\| + \|l(\bar{x}, \theta_0)\|) \} (\|x - \bar{x}\| + \|y - \bar{y}\|) \\
& =: f_{\theta_0}(x, \bar{x}, y, \bar{y}) (\|x - \bar{x}\| + \|y - \bar{y}\|).
\end{aligned}$$

Thus the prerequisites of Theorem 2.2 of [25] are satisfied and the rest of the proof can be carried out in complete analogy to the proof of Theorem 2.1 above. \square

Proof of Lemma 3.1. It is sufficient to verify the existence of some $\eta > 0$ with $P(n^{-1}\widehat{G}_n > \eta) \rightarrow_{n \rightarrow \infty} 1$. According to the continuity of characteristic functions, there exist $-\infty < T_{1,i} < T_{2,i} < \infty$, $i = 1, \dots, d$, such that

$$|c_X(t) - c(t, \theta_0)| > c_0, \quad \forall t \in \Omega_T := [T_{1,1}, T_{2,1}] \times \dots \times [T_{1,d}, T_{2,d}],$$

for some $c_0 > 0$. This implies

$$\begin{aligned}
P(n^{-1}\widehat{G}_n > \eta) & \geq P \left(\int_{\Omega_T} \left| c_n(t) - c_X(t) + c_X(t) - c(t, \theta_0) - (\widehat{\theta}_n - \theta_0)' c^{(1)}(t, \theta_0) \right. \right. \\
& \quad \left. \left. - \frac{1}{2} (\widehat{\theta}_n - \theta_0)' c^{(2)}(t, \widetilde{\theta}) (\widehat{\theta}_n - \theta_0) \right|^2 w(t) dt > \eta \right) \\
& \geq P \left(\int_{\Omega_T} |c_X(t) - c(t, \theta_0)|^2 w(t) dt > 2\eta, \right. \\
& \quad \left. 4 \int_{\Omega_T} \left\{ |c_n(t) - c_X(t)| + \left| (\widehat{\theta}_n - \theta_0)' c^{(1)}(t, \theta_0) + \frac{1}{2} (\widehat{\theta}_n - \theta_0)' c^{(2)}(t, \widetilde{\theta}) (\widehat{\theta}_n - \theta_0) \right| \right\} w(t) dt \leq \eta \right)
\end{aligned}$$

for some random $\tilde{\theta}$ between $\hat{\theta}_n$ and θ_0 . For sufficiently small $\eta > 0$ we obtain

$$\begin{aligned} P(n^{-1}\hat{G}_n > \eta) &\geq P\left(\int_{\Omega_T} \left\{ |c_n(t) - c_X(t)| + \left| (\hat{\theta}_n - \theta_0)' c^{(1)}(t, \theta_0) + \frac{1}{2} (\hat{\theta}_n - \theta_0)' c^{(2)}(t, \tilde{\theta}) (\hat{\theta}_n - \theta_0) \right| \right\} w(t) dt \leq \eta/4 \right) \\ &\geq P\left(\int_{\Omega_T} \left| (\hat{\theta}_n - \theta_0)' c^{(1)}(t, \theta_0) + \frac{1}{2} (\hat{\theta}_n - \theta_0)' c^{(2)}(t, \tilde{\theta}) (\hat{\theta}_n - \theta_0) \right| w(t) dt \leq \eta/8 \right) \\ &\quad + P\left(\int_{\Omega_T} |c_n(t) - c_X(t)| w(t) dt \leq \eta/8 \right) - 1. \end{aligned}$$

The first probability on the r.h.s. tends to one. To show that

$$P\left(\int_{\Omega_T} |c_n(t) - c_X(t)| w(t) dt < \eta/8\right) \xrightarrow{n \rightarrow \infty} 1, \quad (4.7)$$

one can proceed as in the proof of [Lemma 2.2](#). On the one hand,

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P\left(\sup_{\|s-t\| \leq \delta} |c_n(s) - c_X(s) - c_n(t) + c_X(t)| \geq \varepsilon\right)$$

tends to zero according to uniform Lipschitz continuity of $(c_n)_{n \in \mathbb{N}}$ and uniform continuity of c . On the other hand, pointwise convergence $c_n(t) \xrightarrow{P} c_X(t)$, $t \in \mathbb{R}^d$, follows from inequality (4.5) in the proof of [Theorem 3.1](#). \square

Proof of Proposition 3.1. This is an immediate consequence of [Proposition 4.1](#) and [Lemma 4.3](#). \square

Proof of Proposition 3.2. Similarly to the previous proof, the bootstrap statistic can be approximated as follows:

$$\begin{aligned} \hat{G}_n^* &= n \int_{\mathbb{R}^d} |c_n^*(t) - c^*(t) - (\hat{\theta}_n^* - \hat{\theta}_n)' c^{(1)}(t, \hat{\theta}_n)|^2 w(t) dt + o_{P^*}(1) \\ &= n \int_{\mathbb{R}^d} \left| c_n^*(t) - c^*(t) - \left(\frac{1}{n} \sum_{j=1}^n l(X_j^*, \hat{\theta}_n) \right)' c^{(1)}(t, \hat{\theta}_n) \right|^2 w(t) dt + o_{P^*}(1). \end{aligned}$$

Thus, it suffices to show that $G_n^* \xrightarrow{d} Z_2$ in probability, where

$$G_n^* = n \int_{\mathbb{R}^d} \left| c_n^*(t) - c^*(t) - \left(\frac{1}{n} \sum_{j=1}^n l(X_j^*, \hat{\theta}_n) \right)' c^{(1)}(t, \hat{\theta}_n) \right|^2 w(t) dt.$$

The direct application of the result of [\[25\]](#) is again not possible since G_n^* is not the bootstrap counterpart of G_n , but

$$\begin{aligned} \tilde{V}_n^* &= \frac{1}{n} \sum_{j,k=1}^n \left[h_{\hat{\theta}_n}(X_j^*, X_k^*) + \iint_{\mathbb{R}^d \times \mathbb{R}^d} h_{\hat{\theta}_n}(x, y) P_{X_1^*|X_1, \dots, X_n}(dx) P_{X_1^*|X_1, \dots, X_n}(dy) \right. \\ &\quad \left. - \int_{\mathbb{R}^d} h_{\hat{\theta}_n}(x, X_k^*) P_{X_1^*|X_1, \dots, X_n}(dx) - \int_{\mathbb{R}^d} h_{\hat{\theta}_n}(X_j^*, y) P_{X_1^*|X_1, \dots, X_n}(dy) \right] \\ &\xrightarrow{d} Z_2 \quad \text{in probability,} \end{aligned}$$

where $h_{\hat{\theta}_n}$ is defined by a substitution of θ_0 through $\hat{\theta}_n$ in (4.6). Moreover, straightforward calculations yield that $G_n^* = \tilde{V}_n^*$. That is, with probability tending to one, G_n^* is the artificially degenerated bootstrap counterpart of G_n . Therefore, $G_n^* \xrightarrow{d} Z_2$ in probability, which also implies the second assertion of the proposition due to the continuity of the distribution function of Z_2 in case of a positive variance. \square

Proof of Corollary 3.1. Step 1. $\lim_{n \rightarrow \infty} P(\hat{G}_n > t_\alpha^*) = \alpha$.

This assertion follows from [Proposition 3.2](#).

Step 2. $\lim_{n \rightarrow \infty} P(\hat{G}_n > t_\alpha^*) = 1$.

Note that $\hat{G}_n^* - G_n^* = o_{P^*}(1)$ holds under \mathcal{H}_1 as well and that G_n^* is still degenerate. Thus, the bootstrap quantiles are bounded in probability. Now, the claim follows immediately from [Lemma 3.1](#).

Step 3. $\liminf_{n \rightarrow \infty} P_{\mathcal{H}_{1,n}}(\hat{G}_n > t_\alpha^*) \geq \alpha$.

With the same arguments as in the proof of [Corollary 2.1](#), this inequality can be deduced from [Proposition 3.1](#). \square

4.2. Auxiliary results on V -statistics

First we state an auxiliary result concerning variances of partial sums of functions of τ -dependent observations which is extensively applied within the proofs of our main results.

Lemma 4.1. Let $(X_{n,k})_{k=1}^n$, $n \in \mathbb{N}$, be a triangular scheme of \mathbb{R}^d -valued random variables such that (S2)(i) holds for some $\delta \in (0, 1)$. Suppose that the sequence of functions $g_n: \mathbb{R}^d \rightarrow \mathbb{R}$ satisfies $\sup_{n \in \mathbb{N}} \mathbb{E}|g_n(X_{n,1})|^{(2-\delta)/(1-\delta)} < \infty$ and

$$|g_n(x) - g_n(y)| \leq f_n(x, y)\|x - y\|, \quad \forall x, y \in \mathbb{R}^d,$$

where the functions $f_n: \mathbb{R}^{2d} \rightarrow \mathbb{R}$ fulfil $\sup_{n \in \mathbb{N}} \mathbb{E}\{[f_n(X, Y)]^{1/(1-\delta)}(\|X\| + \|Y\|)\} < \infty$ for i.i.d. random variables X, Y with $X \sim P_{X_{n,1}}$. Then,

$$\sup_{n \in \mathbb{N}} \text{var} \left[\frac{1}{\sqrt{n}} \sum_{k=1}^n g_n(X_{n,k}) \right] < \infty.$$

Proof. We have

$$\text{var} \left[\frac{1}{\sqrt{n}} \sum_{k=1}^n g_n(X_{n,k}) \right] = \frac{1}{n} \sum_{k=1}^n \text{var}(g_n(X_{n,k})) + \frac{2}{n} \sum_{1 \leq j < k \leq n} \text{cov}(g_n(X_{n,j}), g_n(X_{n,k})).$$

Denote by $\tilde{X}_{n,k}$ a copy of $X_{n,k}$ that is independent of $X_{n,j}$ and that satisfies $\mathbb{E}\|X_{n,k} - \tilde{X}_{n,k}\| \leq \bar{\tau}_{k-j}$. According to (1.1) such a random variable exists, at least after enlarging the underlying probability space. Based on this relation we can bound the covariance terms by an iterative application of Hölder's inequality,

$$\begin{aligned} |\text{cov}(g_n(X_{n,j}), g_n(X_{n,k}))| &= |\mathbb{E}g_n(X_{n,j})[g_n(X_{n,k}) - g_n(\tilde{X}_{n,k})]| \\ &\leq (\mathbb{E}|g_n(X_{n,j})|^{1/(1-\delta)} |g_n(X_{n,k}) - g_n(\tilde{X}_{n,k})|)^{1-\delta} (\mathbb{E}|g_n(X_{n,k}) - g_n(\tilde{X}_{n,k})|)^\delta \\ &\leq C(\mathbb{E}[f_n(X_{n,k}, \tilde{X}_{n,k})]^{1/(1-\delta)} (\|X_{n,k}\| + \|\tilde{X}_{n,k}\|))^{\delta(1-\delta)} (\bar{\tau}_{k-j})^{\delta^2}. \end{aligned}$$

Eventually, this leads to

$$\text{var} \left[\frac{1}{\sqrt{n}} \sum_{k=1}^n g_n(X_{n,k}) \right] \leq C + \frac{C}{n} \sum_{1 \leq j < k \leq n} (\bar{\tau}_{k-j})^{\delta^2} \leq C \left[1 + \sum_{r=1}^{\infty} (\bar{\tau}_r)^{\delta^2} \right] < \infty,$$

which in turn implies the assertion. \square

Next, we devise a general result on V -statistics for triangular schemes of random variables. Besides (S2) we assume

- (A1) (i) The kernel $h: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a symmetric, measurable function and degenerate under P_{X_1} , i.e. $\mathbb{E}h(x; X_1) = 0$, $\forall x \in \mathbb{R}^d$.
 (ii) For a δ satisfying (S2), the following moment constraints hold true with some $\nu > (2 - \delta)/(1 - \delta)$ and an independent copy \tilde{X}_1 of X_1 :

$$\sup_{1 \leq k < n, n \in \mathbb{N}} \mathbb{E}|h(X_{n,1}, X_{n,1+k})|^\nu < \infty \quad \text{and} \quad \mathbb{E}|h(X_1, \tilde{X}_1)|^\nu + \mathbb{E}|h(X_1, X_1)| < \infty.$$

- (iii) With a continuous function $f: \mathbb{R}^{4d} \rightarrow \mathbb{R}$, the kernel satisfies

$$|h(x, y) - h(\bar{x}, \bar{y})| \leq f(x, \bar{x}, y, \bar{y})[\|x - \bar{x}\| + \|y - \bar{y}\|], \quad \forall x, \bar{x}, y, \bar{y} \in \mathbb{R}^d.$$

Moreover, let $\delta \in (0, 1)$ be such that (A1)(ii) holds. Then

$$\sup \mathbb{E} \left(\max_{a_1, a_2 \in [-A, A]^d} [f(Y_{n,1}, Y_{n,2} + a_1, Y_{n,3}, Y_{n,4} + a_2)]^{1/(1-\delta)} \|Y_{n,5}\| \right) < \infty$$

for some $A > 0$. Here the supremum is taken over all $Y_{n,k}$, $n \in \mathbb{N}$, where for every $k \in \{1, \dots, 5\}$ either $Y_{n,k} \sim P_{X_1}$ or $Y_{n,k} \sim P_{X_{n,1}}$.

This assumption can be seen as the counterpart of (A2) and (A4) of [25], who considered stationary sequences of random variables, for triangular schemes of random variables. In particular, note that any kernel H that is degenerate under P_{X_1} is asymptotically degenerate under $P_{X_{n,1}}$ in the sense of

$$\left| \int_{\mathbb{R}^d} H(x, y) P_{X_{n,1}}(dx) P_{X_{n,1}}(dy) \right| \leq \frac{1}{n} \|g_n\|_\infty^2 \mathbb{E}|H(X_1, \tilde{X}_1)| \xrightarrow{n \rightarrow \infty} 0.$$

This turns out to be crucial for proving the asymptotic result below.

Proposition 4.1. Suppose that the assumptions (S2) and (A1) are fulfilled. Then,

$$nV_{n,n} := \frac{1}{n} \sum_{j,k=1}^n h(X_{n,j}, X_{n,k}) \xrightarrow{d} Z_{loc},$$

$$Z_{loc} := \lim_{c \rightarrow \infty} \sum_{j=0}^{\infty} \sum_{k_1, k_2 \in \mathbb{Z}^d} \gamma_{j; k_1, k_2}^{(c)} [(Z_{j; k_1} + D_{j; k_1})(Z_{j; k_2} + D_{j; k_2}) - C_{j; k_1, k_2}] + \mathbb{E}h(X_1, X_1), \quad (4.8)$$

where the r.h.s. converges in the L_2 -sense for certain families of constants $(\gamma_{j; k_1, k_2}^{(c)})_{c, j, k_1, k_2}$, $(C_{j; k_1, k_2})_{j, k_1, k_2}$ and $(D_{j; k})_{j, k}$ and centred, jointly normal random variables $(Z_{j; k})_{j, k}$.

Proof. According to the law of large numbers (Lemma 5.1 of [25]), we obtain

$$\frac{1}{n} \sum_{k=1}^n h(X_{n,k}, X_{n,k}) \xrightarrow{P} \mathbb{E}h(X_1, X_1).$$

Thus it remains to investigate the limit of the U -statistic

$$nU_{n,n} := \frac{1}{n} \sum_{1 \leq j \neq k \leq n} h(X_{n,j}, X_{n,k}).$$

The main ideas of the proof are already contained in the proof of Theorem 2.2 in [25]. Therefore we only point out the additional calculations that are required for the extension of her results on sequences of random variables to triangular schemes. In a first step one reduces the problem to statistics with bounded kernels. To this end, we introduce degenerate kernels

$$h_c(x, y) = \tilde{h}_c(x, y) - \int_{\mathbb{R}^d} \tilde{h}_c(x, y) P_{X_1}(dx) - \int_{\mathbb{R}^d} \tilde{h}_c(x, y) P_{X_1}(dy) + \iint_{\mathbb{R}^d \times \mathbb{R}^d} \tilde{h}_c(x, y) P_{X_1}(dx) P_{X_1}(dy)$$

such that

- (1) $\tilde{h}_c(x, y) = h(x, y)$, $\forall x, y \in [-c, c]$,
- (2) \tilde{h}_c is bounded and $(h_c)_c$ satisfies (A1)(iii) uniformly,
- (3) $\sup_{c \in \mathbb{R}_+} (\sup_{k \in \mathbb{N}} \mathbb{E}|h_c(X_{n,1}, X_{n,1+k})|^v + \mathbb{E}|h_c(X_1, \tilde{X}_1)|^v + \mathbb{E}|h_c(X_1, X_1)|) < \infty$,

e.g. $\tilde{h}_c = (h \wedge \max_{x, y \in [-c, c]^d} |h(x, y)|) \vee (-\max_{x, y \in [-c, c]^d} |h(x, y)|)$. Denote the corresponding U -statistic by $U_{n,n,c}$. In order to show that $\limsup_{n \rightarrow \infty} n^2 \mathbb{E}(U_{n,n} - U_{n,n,c})^2 \rightarrow 0$ as $c \rightarrow \infty$, first note that $U_{n,n} - U_{n,n,c}$ is a U -statistic with kernel $H = h - h_c$. The second moment of this statistic can be bounded using the decomposition

$$n^2 \mathbb{E}(U_{n,n} - U_{n,n,c})^2 \leq 8 \sup_{1 \leq k < n} \mathbb{E}|H(X_{n,1}, X_{n,1+k})|^2 + \frac{8}{n} \sum_{r=1}^{n-1} \sum_{t=1}^4 Z_{n,r}^{(t)}$$

with

$$Z_{n,r}^{(1)} := \sum_{\substack{1 \leq i < j; k < l; j \leq l \leq n \\ r := \min\{j, k\} - i \geq l - \max\{j, k\}}} \{|\mathbb{E}H(X_{n,i}, \tilde{X}_{n,j}^{(r)})H(\tilde{X}_{n,k}^{(r)}, \tilde{X}_{n,l}^{(r)})| + |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, X_{n,l}) - \mathbb{E}H(X_{n,i}, \tilde{X}_{n,j}^{(r)})H(\tilde{X}_{n,k}^{(r)}, \tilde{X}_{n,l}^{(r)})|\},$$

$$Z_{n,r}^{(2)} := \sum_{\substack{1 \leq i < j; i \leq k; k < l \leq n \\ r := l - \max\{j, k\} > \min\{j, k\} - i}} \{|\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, \tilde{X}_{n,l}^{(r)})| + |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, X_{n,l}) - \mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, \tilde{X}_{n,l}^{(r)})|\},$$

$$Z_{n,r}^{(3)} := \sum_{\substack{1 \leq i \leq k < l < j \leq n \\ r := k - i \geq j - l}} \{|\mathbb{E}H(X_{n,i}, \tilde{X}_{n,j}^{(r)})H(\tilde{X}_{n,k}^{(r)}, \tilde{X}_{n,l}^{(r)})| + |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, X_{n,l}) - \mathbb{E}H(X_{n,i}, \tilde{X}_{n,j}^{(r)})H(\tilde{X}_{n,k}^{(r)}, \tilde{X}_{n,l}^{(r)})|\},$$

$$Z_{n,r}^{(4)} := \sum_{\substack{1 \leq i \leq k < l < j \leq n \\ r := j - l > k - i}} \{|\mathbb{E}H(X_{n,i}, \tilde{X}_{n,j}^{(r)})H(X_{n,k}, X_{n,l})| + |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, X_{n,l}) - \mathbb{E}H(X_{n,i}, \tilde{X}_{n,j}^{(r)})H(X_{n,k}, X_{n,l})|\}.$$

In every summand of $Z_{n,r}^{(1)}$ and $Z_{n,r}^{(3)}$ the vector $(\tilde{X}_{n,j}^{(r)}, \tilde{X}_{n,k}^{(r)}, \tilde{X}_{n,l}^{(r)})'$ is chosen such that it is independent of the random variable $X_{n,i}$, $(\tilde{X}_{n,j}^{(r)}, \tilde{X}_{n,k}^{(r)}, \tilde{X}_{n,l}^{(r)})' \stackrel{d}{=} (X'_{n,j}, X'_{n,k}, X'_{n,l})'$, and (1.1) holds. Within $Z_{n,r}^{(2)}$ (respectively $Z_{n,r}^{(4)}$) the random variable $\tilde{X}_{n,l}^{(r)}$ (respectively $\tilde{X}_{n,j}^{(r)}$) is chosen to be independent of the vector $(X'_{n,i}, X'_{n,j}, X'_{n,k})'$ (respectively $(X'_{n,i}, X'_{n,k}, X'_{n,l})'$) such that $\tilde{X}_{n,l}^{(r)} \stackrel{d}{=} X_{n,l}$ (respectively $\tilde{X}_{n,j}^{(r)} \stackrel{d}{=} X_{n,j}$) and (1.1) holds. This may possibly require an enlargement of the underlying probability space. Moreover, note that the number of summands of $Z_{n,r}^{(t)}$, $t = 1, \dots, 4$, is bounded by $(r+1)n^2$. Asymptotic negligibility

of the second summands of $Z_{n,r}^{(t)}$ can be verified along the lines of the proof of Lemma 3.1 of [25]. The remaining terms, that are equal to zero her setting, can be bounded with the help of L_1 -coupling techniques, the asymptotic degeneracy of H and the properties of the truncated kernels h_c . Exemplarily we consider the first summands of $Z_{n,2}$ which can be estimated from above by

$$\begin{aligned} & \frac{1}{n^2} \sum_{r=1}^{n-1} \sum_{1 \leq i < j \leq k < l \leq n; r:=l-k > j-i; k-j \geq j-i} |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, \tilde{X}_{n,l}^{(r)})| \\ & + \frac{1}{n^2} \sum_{r=1}^{n-1} \sum_{1 \leq i < j \leq k < l \leq n; r:=l-k > j-i > k-j} |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, \tilde{X}_{n,l}^{(r)})| \\ & + \frac{1}{n^2} \sum_{r=1}^{n-1} \sum_{1 \leq i \leq k < j < l \leq n; r:=l-j > k-i; j-k \geq k-i} |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, \tilde{X}_{n,l}^{(r)})| \\ & + \frac{1}{n^2} \sum_{r=1}^{n-1} \sum_{1 \leq i \leq k < j < l \leq n; r:=l-j > k-i > j-k} |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, \tilde{X}_{n,l}^{(r)})|. \end{aligned}$$

We investigate the first summand only since the remaining terms can be treated in an analogous manner. Let $\tilde{X}_{n,k} \sim P_{X_{n,k}}$ be independent of the vector $(X'_{n,i}, X'_{n,j})'$ and such that $\mathbb{E}|\tilde{X}_{n,k} - X_{n,k}| \leq \bar{\tau}_{k-j}$. (This may require an enlargement of the underlying probability space.) Now an iterative application of Hölder's inequality yields

$$\begin{aligned} |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, \tilde{X}_{n,l}^{(r)})| & \leq \left| \mathbb{E}H(X_{n,i}, X_{n,j}) \left[\int_{\mathbb{R}^d} H(X_{n,k}, y) - H(\tilde{X}_{n,k}, y) P_{X_{n,1}}(dy) \right] \right| + |\mathbb{E}H(X_{n,i}, X_{n,j})\mathbb{E}H(X_{n,k}, \tilde{X}_{n,l}^{(r)})| \\ & \leq C \bar{\tau}_{k-j}^2 (\mathbb{E}|H(X_{n,i}, X_{n,j})|^{(2-\delta)/(1-\delta)})^{(1-\delta)/(2-\delta)} + |\mathbb{E}H(X_{n,i}, X_{n,j})| \iint_{\mathbb{R}^d \times \mathbb{R}^d} |H(x, y)| P_{X_1}(dx) P_{X_1}(dy) \frac{\|g_n\|_\infty}{n} \\ & \leq C (\mathbb{E}|H(X_{n,i}, X_{n,j})|^{(2-\delta)/(1-\delta)})^{(1-\delta)/(2-\delta)} \left[\bar{\tau}_{k-j}^2 + \frac{\|g_n\|_\infty}{n} \bar{\tau}_{j-i}^\delta \right] + C \left[\iint_{\mathbb{R}^d \times \mathbb{R}^d} |H(x, y)| P_{X_1}(dx) P_{X_1}(dy) \frac{\|g_n\|_\infty}{n} \right]^2. \end{aligned}$$

Within these calculations, the inequality

$$\begin{aligned} \mathbb{E} \int_{\mathbb{R}^d} |H(X_{n,k}, y) - H(\tilde{X}_{n,k}, y)| P_{X_{n,1}}(dy) & \leq \bar{\tau}_{k-j}^\delta \left[\iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} [f_1(x, z, y, y)]^\eta (\|x\| + \|z\|) P_{X_{n,1}}(dx) \right. \\ & \quad \times (dy) P_{X_{n,1}}(dz) \left. \right]^{1-\delta} \\ & \leq C \bar{\tau}_{k-j}^\delta \end{aligned}$$

is applied which is a consequence of (A1)(iii), where the function f_1 is given by

$$f_1(x, \bar{x}, y, \bar{y}) := 2f(x, \bar{x}, y, \bar{y}) + \int_{\mathbb{R}^d} f(x, \bar{x}, y, y) P_{X_1}(dy) + \int_{\mathbb{R}^d} f(x, x, y, \bar{y}) P_{X_1}(dx).$$

Similarly to the investigation of the term E_1 in the proof of Lemma 2.1 in [25], it can be shown that there exists a family $(\varepsilon_c)_{c \in \mathbb{R}_+}$ with $\varepsilon_c \rightarrow_{c \rightarrow \infty} 0$ such that

$$\left(\sup_{1 \leq k < n; n \in \mathbb{N}} \mathbb{E}|H(X_{n,1}, X_{n,1+k})|^{(2-\delta)/(1-\delta)} \right)^{(1-\delta)/(2-\delta)} + \mathbb{E}|H(X_1, \tilde{X}_1)| \leq \varepsilon_c.$$

Here, \tilde{X}_1 denotes an independent copy of X_1 . Hence, the relation

$$\limsup_{n \rightarrow \infty} \frac{1}{n^2} \sum_{r=1}^{n-1} \sum_{1 \leq i < j \leq k < l \leq n; r:=l-k > j-i; k-j \geq j-i} |\mathbb{E}H(X_{n,i}, X_{n,j})H(X_{n,k}, \tilde{X}_{n,l}^{(r)})| \leq C \varepsilon_c \xrightarrow{c \rightarrow \infty} 0$$

holds. After similar estimations for the remaining terms one obtains asymptotic negligibility of the error $n^2 \mathbb{E}(U_{n,n} - U_{n,n,c})^2$. In a next step, the statistics with bounded kernels are approximated by their (finite) multi-scale wavelet series expansion. Similar arguments as in the respective estimation steps in the proof of Theorem 2.2 in [25] lead to

$$\lim_{K \rightarrow \infty} \lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} n^2 \mathbb{E}(U_{n,n,c} - U_{n,n,c}^{(K,L)})^2 = 0.$$

Here,

$$U_{n,n,c}^{(K,L)} = \sum_{j=0}^{J(K)} \sum_{k_1, k_2 \in \{-L, \dots, L\}^d} \gamma_{j;k_1, k_2}^{(c)} \left\{ \frac{1}{n} \sum_{s,t=1}^n [\psi_{j;k_1}(X_{n,s}) - \mathbb{E}\psi_{j;k_1}(X_1)][\psi_{j;k_2}(X_{n,t}) - \mathbb{E}\psi_{j;k_2}(X_1)] \right. \\ \left. - \frac{1}{n} \sum_{t=1}^n [\psi_{j;k_1}(X_{n,t}) - \mathbb{E}\psi_{j;k_1}(X_1)][\psi_{j;k_2}(X_{n,t}) - \mathbb{E}\psi_{j;k_2}(X_1)] \right\} \quad (4.9)$$

with $\psi_{j;k} := \Phi_{j;l/2}^{(e)}$, $l + e = k$, $l/2 \in \mathbb{Z}^d$ and $e \in \{0, 1\}^d$. The functions $\Phi_{j;k}^{(e)}$ and the constants $\gamma_{j;k_1, k_2}^{(c)}$ are defined as in the proof of Theorem 2.1 (with h_c instead of $h_{\mu_0}^{(c)}$). The subtrahend of the r.h.s. of (4.9) converges in probability to $C_{j;k_1, k_2}$ due to the weak law of large numbers, where $(C_{j;k_1, k_2})_{j, k_1, k_2}$ is defined as in the proof of Theorem 2.1. Moreover, we get

$$\sqrt{n}(\mathbb{E}\psi_{j;k}(X_{n,1}) - \mathbb{E}\psi_{j;k}(X_1)) \xrightarrow{P} D_{j;k} := \int_{\mathbb{R}^d} [\psi_{j;k}(x) - \mathbb{E}\psi_{j;k}(X_1)]g(x)P_{X_1}(dx).$$

In conjunction with a CLT, Slutsky's Lemma, and the continuous mapping, this yields

$$U_{n,n,c}^{(K,L)} \xrightarrow{d} Z_c^{(K,L)}$$

with

$$Z_c^{(K,L)} := \sum_{j=0}^{J(K)} \sum_{k_1, k_2 \in \{-L, \dots, L\}^d} \gamma_{j;k_1, k_2}^{(c)} [(Z_{j;k_1} + D_{j;k_1})(Z_{j;k_2} + D_{j;k_2}) - C_{j;k_1, k_2}],$$

where $(Z_{j;k})_{j,k}$ is defined as in the proof of Theorem 2.1. Here, a slight modification of the CLT of [29] can be employed; see the proof of Theorem 3.1 in [25] for details. Since additionally

$$\lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} \lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} n^2 \mathbb{E}(U_{n,n} - U_{n,n,c}^{(K,L)})^2.$$

Theorem 3.2 of [6] implies that it remains to show

$$\lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} \lim_{L \rightarrow \infty} \mathbb{E}(Z_c^{(K,L)} - Z_{loc} + \mathbb{E}h(X_1, X_1))^2 = 0.$$

This in turn follows from the corresponding relation of the U -statistics by an iterative application of Theorem 3.4 of [6]. \square

According to [25], the conditions (S2), (A1) and $\sup_{k \in \mathbb{N}} \mathbb{E}|h(X_1, X_{k+1})|^v < \infty$ are sufficient for

$$nV_n = \frac{1}{n} \sum_{j,k=1}^n h(X_j, X_k) \xrightarrow{d} Z$$

with

$$Z := \lim_{c \rightarrow \infty} \sum_{j=0}^{\infty} \sum_{k_1, k_2 \in \mathbb{Z}^d} \gamma_{j;k_1, k_2}^{(c)} [Z_{j;k_1} Z_{j;k_2} - C_{j;k_1, k_2}] + \mathbb{E}h(X_1, X_1). \quad (4.10)$$

Below we compare the asymptotic behaviour of nV_n and $nV_{n,n}$ when the kernels are of the form

$$h(x, y) = \int_{\mathbb{R}^d} [g(x, t)]' g(y, t) w(t) dt \quad (4.11)$$

with some vector-valued function g .

Lemma 4.2. Suppose that the conditions (S2) and (A1) are satisfied. Moreover, let the kernel h be of the form (4.11) with $\mathbb{E} \int_{\mathbb{R}^d} \|g(X_1, t)\|^{2v} w(t) dt < \infty$ and $\text{var}(Z) > 0$. Then,

$$P(Z_{loc} > x) \geq P(Z > x), \quad \forall x \in \mathbb{R},$$

where Z and Z_{loc} are defined as in (4.10) and Proposition 4.1, respectively.

Proof. First note that the limits of V_n and $V_{n,n}$ can alternatively be expressed using a uni-scale wavelet decomposition, i.e.

$$Z \stackrel{d}{=} \lim_{c \rightarrow \infty} \lim_{J \rightarrow \infty} \sum_{k_1, k_2 \in \mathbb{Z}^d} \alpha_{j;k_1, k_2}^{(c)} [Z_{j;k_1} Z_{j;k_2} - A_{j;k_1, k_2}]$$

and

$$Z_{loc} \stackrel{d}{=} \lim_{c \rightarrow \infty} \lim_{J \rightarrow \infty} \sum_{k_1, k_2 \in \mathbb{Z}^d} \alpha_{j; k_1, k_2}^{(c)} [(Z_{j, k_1} + C_{j; k_1})(Z_{j, k_2} + C_{j; k_2}) - A_{j; k_1, k_2}],$$

where the r.h.s. converge in the L_2 -sense. Here, $(Z_{j, k})_{j, k}$ are centred and jointly normal random variables with covariance structure

$$\text{cov}(Z_{j, s}, Z_{j, t}) = \text{cov}(\Phi_{j; s}^{(0d)}(X_1), \Phi_{j; t}^{(0d)}(X_1)) + \sum_{k=2}^{\infty} [\text{cov}(\Phi_{j; s}^{(0d)}(X_1), \Phi_{j; t}^{(0d)}(X_k)) + \text{cov}(\Phi_{j; s}^{(0d)}(X_k), \Phi_{j; t}^{(0d)}(X_1))].$$

$C_{j; k} := \int_{\mathbb{R}^d} [\Phi_{j; k}(x) - \mathbb{E}\Phi_{j; k}(X_1)]g(x)P_{X_1}(dx)$, $\alpha_{j; k_1, k_2}^{(c)} := \int_{\mathbb{R}^d \times \mathbb{R}^d} h_c(x, y)\Phi_{j; k_2}^{(0d)}(x)\Phi_{j; k_1}^{(0d)}(y)dx dy$, and $A_{j; k_1, k_2} := \text{cov}(\Phi_{j, k_1}^{(0d)}(X_1), \Phi_{j, k_2}^{(0d)}(X_1))$. To define h_c , we first introduce $G_c(t) := \max_{x \in [-c, c]^d} |g(x, t)|$ and truncated versions of g ,

$$g^{(c)}(x, t) := \begin{cases} g(x, t) & \text{for } |g(x, t)| \leq G_c(t), \\ -G_c(t) & \text{for } g(x, t) < -G_c(t), \\ G_c(t) & \text{for } g(x, t) > G_c(t). \end{cases}$$

The truncated versions $(h_c)_c$ of the kernel h itself are then defined through

$$h_c(x, y) := \int_{\mathbb{R}^d} \left[g^{(c)}(x, t) - \int_{\mathbb{R}^d} g^{(c)}(z, t)P_{X_1}(dz) \right] \left[g^{(c)}(y, t) - \int_{\mathbb{R}^d} g^{(c)}(z, t)P_{X_1}(dz) \right] w(t) dt.$$

We show that for all $\varepsilon > 0$ and $x \in \mathbb{R}$ the inequality $P(Z_{loc} > x) - P(Z > x) \geq -\varepsilon$ holds true. W.l.o.g. x can assumed to be a continuity point of the cumulative distribution function of $Z_{loc} - \mathbb{E}h(X_1, X_1)$. It suffices to verify

$$P(\tilde{Z}_{loc, c}^{(J, L)} > x) - P(\tilde{Z}_c^{(J, L)} > x) \geq 0, \quad \forall c \geq c_0, L \geq L_0(c), J \geq J_0(c, L),$$

where

$$\tilde{Z}_{loc, c}^{(J, L)} := \sum_{k_1, k_2 \in \{-L, \dots, L\}^d} \alpha_{j; k_1, k_2}^{(c)} [(Z_{j, k_1} + C_{j; k_1})(Z_{j, k_2} + C_{j; k_2}) - A_{j; k_1, k_2}]$$

and

$$\tilde{Z}_c^{(J, L)} := \sum_{k_1, k_2 \in \{-L, \dots, L\}^d} \alpha_{j; k_1, k_2}^{(c)} [Z_{j, k_1} Z_{j, k_2} - A_{j; k_1, k_2}].$$

Let the constants J, L and c be fixed. Then, with $y = x + B_c^{(J, L)}$ and $B_c^{(J, L)} := \sum_{k_1, k_2 \in \{-L, \dots, L\}^d} \alpha_{j; k_1, k_2}^{(c)} A_{j; k_1, k_2}$, we have to prove

$$P([Z^{(J, L)} + C^{(J, L)}]' \Delta_c^{(J, L)} [Z^{(J, L)} + C^{(J, L)}] > y) \geq P([Z^{(J, L)}]' \Delta_c^{(J, L)} Z^{(J, L)} > y). \quad (4.12)$$

Here, $Z^{(J, L)} := (Z_{j, -L1_d}, \dots, Z_{j, L1_d})'$, $C^{(J, L)} := (C_{j, -L1_d}, \dots, C_{j, L1_d})'$ and $\Delta_c^{(J, L)}$ is a symmetric matrix of the corresponding coefficients $\alpha_{j; k_1, k_2}^{(c)}$. The latter inequality is equivalent to

$$P([S_{(J, L)}^{1/2} Y + O'_{(J, L)} C^{(J, L)}]' O'_{(J, L)} \Delta_c^{(J, L)} O_{(J, L)} [S_{(J, L)}^{1/2} Y + O'_{(J, L)} C^{(J, L)}] > y) \geq P([S_{(J, L)}^{1/2} Y]' O'_{(J, L)} \Delta_c^{(J, L)} O_{(J, L)} [S_{(J, L)}^{1/2} Y] > y)$$

for some a diagonal matrix $S_{(J, L)}^{1/2}$ with decreasing diagonal elements $\sigma_1 \geq \dots \geq \sigma_N > \sigma_{N+1} = \dots = \sigma_{(2L+1)^d} = 0$ for an $N \leq (2L+1)^d$, an orthogonal matrix $O_{(J, L)}$, and a standard normal vector Y . Note that $N \geq 1$ if the indices c, J, L are chosen sufficiently large. Moreover, we have

$$z' \Delta_c^{(J, L)} z = \int_{\mathbb{R}^d} \left(\sum_{k=1}^{(2L+1)^d} z_k \int_{\mathbb{R}^d} \Phi_{j, k}(x) [g^{(c)}(x, t) - \mathbb{E}g^{(c)}(X_1, t)] dx \right)^2 w(t) dt \geq 0$$

for arbitrary $z = (z_1, \dots, z_{(2L+1)^d})' \in \mathbb{R}^{(2L+1)^d}$, i.e. $\Delta_c^{(J, L)}$ is positive semi-definite. This in turn implies positive semi-definiteness of $O'_{(J, L)} \Delta_c^{(J, L)} O_{(J, L)}$. Therefore,

$$P([S_{(J, L)}^{1/2} Y + O'_{(J, L)} C^{(J, L)}]' O'_{(J, L)} \Delta_c^{(J, L)} O_{(J, L)} [S_{(J, L)}^{1/2} Y + O'_{(J, L)} C^{(J, L)}] > y) = P \left(\sum_{k=1}^{(2L+1)^d} \lambda_k (Y_k + \delta_k)^2 + \sum_{k=1}^{(2L+1)^d} \eta_k > y \right) \quad (4.13)$$

with certain nonnegative constants λ_k , real-valued constants δ_k, η_k , $k = 1, \dots, (2L+1)^d$, and $\sum_{k=1}^{(2L+1)^d} \eta_k \geq 0$ while

$$P([Z^{(J,L)}]' \Delta_c^{(J,L)} Z^{(J,L)} > y) = P\left(\sum_{k=1}^{(2L+1)^d} \lambda_k Y_k^2 > y\right). \quad (4.14)$$

Note that for a standard normal random variable Y , the variable $(Y+a)^2$ is stochastically larger than Y^2 for any $a \neq 0$. Finally, since $Y_1, \dots, Y_{(2L+1)^d}$ are independent standard normal variables, the comparison of (4.13) and (4.14) implies (4.12). \square

Combining Proposition 4.1 and Lemma 4.2 yields the following assertion.

Lemma 4.3. *Under the assumptions of Lemma 4.2, the following inequality holds true:*

$$\liminf_{n \rightarrow \infty} P[(nV_{n,n} > x) - P(nV_n > x)] \geq 0, \quad \forall x \in \mathbb{R}.$$

Proof. We show that for each $\varepsilon > 0$

$$\liminf_{n \rightarrow \infty} [P(nV_{n,n} > x) - P(nV_n > x)] \geq -\varepsilon, \quad \forall x \in \mathbb{R}.$$

To this end, let the cumulative distribution function of Z_{loc} be continuous at $x + \delta$ for some $\delta > 0$ that is specified below. We obtain

$$\begin{aligned} \liminf_{n \rightarrow \infty} [P(nV_{n,n} > x) - P(nV_n > x)] &\geq \lim_{n \rightarrow \infty} [P(nV_{n,n} > x + \delta) - P(nV_n > x)] \\ &= [P(Z_{loc} > x + \delta) - P(Z > x)] \\ &\geq [P(Z > x + \delta) - P(Z > x)], \end{aligned}$$

where the latter difference is greater than $-\varepsilon$ for sufficiently small δ . \square

Acknowledgments

The author is grateful to Michael H. Neumann for many fruitful discussions on this work. She also thanks an anonymous referee for valuable comments that led to an improvement of the paper. This research was funded by the German Research Foundation DFG (projects: NE 606/2-1 and NE 606/2-2).

References

- [1] A. Antille, G. Kersting, W. Zucchini, Testing symmetry, *J. Amer. Statist. Assoc.* 77 (1982) 639–646.
- [2] M.A. Arcones, E. Giné, On the bootstrap of U and V statistics, *Ann. Statist.* 20 (1992) 655–674.
- [3] J. Bai, S. Ng, Tests for skewness, kurtosis, and normality for time series data, *J. Bus. Econom. Statist.* 23 (2005) 49–60.
- [4] O.E. Barndorff-Nielsen, Processes of normal inverse Gaussian type, *Finance Stoch.* 2 (1998) 41–68.
- [5] P. Bickel, P. Bühlmann, A new mixing notion and functional central limit theorems for a sieve bootstrap in time series, *Bernoulli* 5 (1999) 413–446.
- [6] P. Billingsley, *Convergence of Probability Measures*, 2nd ed., Wiley, New York, 1999.
- [7] H.W. Block, N.A. Langberg, D.S. Stoffer, Time series models for non-Gaussian processes, in: H.W. Block, A.R. Sampson, T.H. Savits (Eds.), *Topics in Statistical Dependence*, Inst. Math. Statist., Hayward, CA, 1990, pp. 69–83.
- [8] G.E.P. Box, G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, Revised edition, Holden-Day, San Francisco, Calif.–Düsseldorf–Johannesburg, 1976.
- [9] P. Bühlmann, Sieve bootstrap for time series, *Bernoulli* 3 (1997) 123–148.
- [10] Y.-T. Chen, R.Y. Chou, C.M. Kuan, Testing time reversibility without moment restrictions, *J. Econometrics* 95 (2000) 199–218.
- [11] J. Dedecker, F. Merlevède, The conditional central limit theorem in Hilbert spaces, *Stochastic Process. Appl.* 108 (2003) 229–262.
- [12] J. Dedecker, C. Prieur, New dependence coefficients. Examples and applications to statistics, *Probab. Theory Related Fields* 132 (2005) 203–236.
- [13] J. Dedecker, C. Prieur, Couplage pour la distance minimale, *C. R. Acad. Sci. Paris, Ser. I* 338 (2004) 805–808.
- [14] P. Doukhan, M.H. Neumann, The notion of ψ -weak dependence and its applications to bootstrapping time series, *Probab. Surv.* 5 (2008) 146–168.
- [15] Y. Fan, Goodness-of-fit tests based on kernel density estimators with fixed smoothing parameters, *Econometric Theory* 14 (1998) 604–621.
- [16] Y. Fan, A. Ullah, On goodness-of-fit tests for weakly dependent processes using kernel method, *J. Nonparametr. Stat.* 11 (1999) 337–360.
- [17] A. Feuerverger, R.A. Mureika, The empirical characteristic function and its applications, *Ann. Statist.* 5 (1977) 88–97.
- [18] B.K. Ghosh, W.-M. Huang, The power and optimal kernel of the Bickel–Rosenblatt test for goodness of fit, *Ann. Statist.* 19 (1991) 999–1009.
- [19] G.G. Gregory, Large sample theory for U -statistics and tests of fit, *Ann. Statist.* 5 (1977) 110–123.
- [20] N. Henze, B. Klar, S.G. Meintanis, Invariant tests for symmetry about an unspecified point based on the empirical characteristic function, *J. Multivariate Anal.* 87 (2003) 275–297.
- [21] R. Ignaccolo, Goodness-of-fit tests for dependent data, *Nonparametr. Statist.* 16 (2004) 19–38.
- [22] W.-C. Ip, H. Wong, M. Chen, Testing normality for linear $AR(p)$ models, *Comm. Statist. Theory Methods* 33 (2004) 891–908.
- [23] K.K. Jose, T. Lishamol, J. Sreekumar, Autoregressive processes with normal-Laplace marginals, *Statist. Probab. Lett.* 78 (2006) 2456–2462.
- [24] W.-M. Lee, Robust M Tests for Distribution Symmetry with Time Series Data, Working Paper, 2007.
- [25] A. Leucht, Degenerate U - and V -statistics under weak dependence: asymptotic theory and bootstrap consistency, *Bernoulli* (2012) (forthcoming).
- [26] A. Munk, J.-P. Stockis, J. Valeinis, G. Giese, Neyman smooth goodness-of-fit tests for the marginal distribution of dependent data, *Ann. Inst. Statist. Math.* 63 (2011) 939–959.
- [27] G. Neuhaus, Asymptotic properties of the Cramér–von Mises-statistic when parameters are estimated, in: J. Hájek (Ed.), *Proceedings of the Prague Symposium on Asymptotic Statistics*, 1974, pp. 257–297.
- [28] M.H. Neumann, E. Paparoditis, On bootstrapping L_2 -type statistics in density testing, *Statist. Probab. Lett.* 50 (2000) 137–147.

- [29] M.H. Neumann, E. Paparoditis, Goodness-of-fit tests for Markovian time series models: central limit theory and bootstrap approximations, *Bernoulli* 14 (2008) 14–46.
- [30] J. Pemberton, H. Tong, A note on the distributions of non-linear autoregressive stochastic models, *J. Time Ser. Anal.* 2 (1981) 49–52.
- [31] J.B. Ramsey, P. Rothman, Time irreversibility and business cycle asymmetry, *J. Money, Credit, Banking* 28 (1996) 1–21.
- [32] R Development Core Team, R: a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2007, URL <http://www.R-project.org>, ISBN 3-900051-07-0.
- [33] X. Shao, W.B. Wu, Asymptotic spectral theory for nonlinear time series, *Ann. Statist.* 35 (2007) 1773–1801.
- [34] R.G. Staudte, S.J. Sheather, *Robust Estimation and Testing*, John Wiley & Sons, New York, 1990.
- [35] E. Taufer, N.N. Leonenko, Simulation of Lévy-driven Ornstein–Uhlenbeck processes with given marginal distribution, *Comput. Statist. Data Anal.* 53 (2009) 2427–2437.
- [36] N.G. Ushakov, *Selected Topics in Characteristic Functions*, VSP, Utrecht, 1999.